



UNIVERSIDADE FEDERAL DO PIAUÍ
CENTRO DE CIÊNCIAS DA NATUREZA
PÓS-GRADUAÇÃO EM MATEMÁTICA
DOUTORADO EM MATEMÁTICA

**Métodos de Ponto Proximal para Diferença de
Campos Vetoriais Monótonos Maximais em
Variedades de Hadamard**

João Santos Andrade

Teresina - 2022

João Santos Andrade

Tese de Doutorado:

**Métodos de Ponto Proximal para Diferença de Campos
Vetoriais Monótonos Maximais em Variedades de Hadamard**

Tese de Doutorado submetida a coordenação do Programa de Pós-Graduação em Matemática, da Universidade Federal do Piauí, como requisito parcial para obtenção do grau de Doutor em Matemática. Área de Concentração Otimização.

Orientador:

Prof. Dr. Jurandir de Oliveira Lopes

Co-Orientador:

Prof. Dr. João Carlos de Oliveira Souza

Teresina - 2022



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO PIAUÍ
CENTRO DE CIÊNCIAS DA NATUREZA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

*Métodos de Ponto Proximal para Diferença de Campos Vetoriais Monótonos
Maximais em Variedades de Hadamard*

João Santos Andrade


Esta Tese foi submetida como parte dos requisitos necessários à obtenção do grau de **Doutor em Matemática**, outorgado pela Universidade Federal do Piauí.


A citação de qualquer trecho deste trabalho é permitida, desde que seja feita em conformidade com as normas da ética científica.

Tese aprovada em 20 de setembro de 2022.

Banca Examinadora:


Prof. Dr. Jurandir de Oliveira Lopes - Orientador


Prof. Dr. João Carlos de Oliveira Souza - Co-Orientador UFPI


Prof. Dr. João Xavier da Cruz Neto - UFPI


Profa. Dra. Márcia Helena Costa Fampa - UFRJ


Prof. Dr. Jefferson Divino Gonçalves de Melo - UFG

FICHA CATALOGRÁFICA
Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca José Albano de Macêdo

A553m Andrade, João Santos

Métodos de ponto proximal para diferença de campos vetoriais monótonos maximais em variedades de Hadamard / João Santos Andrade -- 2022.

1 CD – ROM : il.; 4 ¼ pol. (92f.)

Tese (Doutorado) - Universidade Federal do Piauí, Centro de Ciências da Natureza, Teresina, 2022.

“Orientador : Dr. Jurandir de Oliveira Lopes”

“Coorientador: Dr. João Carlos de Oliveira Souza”

1. Método de ponto proximal. 2. Diferença de campos vetoriais monótonos maximais. 3. Variedades de Hardamard. I. Lopes, Jurandir de Oliveira. II. Souza, João Carlos de Oliveira. III. Título.

CDD 515

Emanuele Alves Araújo - CRB 3/1290

Dedico este trabalho a meus pais Nivaldo e Maria Francisca, a minha esposa Elcilene e meu filho João Vítor, a meus irmãos e irmãs (Franklivaldo, Liana, Valquíria, Niviane, Vânia, Liziane e Nivaldo Filho), a minha cunhada Nira, a meus avós Raimundo e Angelita (In memoriam).

Agradecimentos

Agradeço primeiramente a Deus pela vida e força para vencer as dificuldades.

Agradeço a minha esposa Elcilene, meu filho João Vítor e minha cunhada Nira por todo apoio, incentivo, atenção e paciência nessa caminhada.

Agradeço aos meus pais, irmãs e irmãos pelo apoio e incentivo nessa conquista.

Agradeço aos meus compadres Aurineide (grande incentivadora) e Jefferson pelo apoio e incentivo durante essa jornada. Em especial agradeço a minha amiga querida Maria de Jesus pela amizade, ajuda e apoio de sempre.

Agradeço aos professores Jurandir de Oliveira Lopes e João Carlos de Oliveira Souza pela escolha do tema, orientação, paciência e compreensão.

Agradeço aos professores João Xavier da Cruz Neto, Márcia Helena Costa Fampa e Jefferson Divino Gonçalves de Melo por aceitarem participar da banca de defesa de Tese de Doutorado e pelas valiosas sugestões e correções do texto, que contribuíram bastante para o enriquecimento deste trabalho.

Agradeço aos professores do curso de Matemática da Universidade Federal do Piauí - UFPI, Campus Senador Helvídio Nunes de Barros - CSHNB, da cidade de Picos por permitirem o meu afastamento para cursar a tão sonhada pós-graduação: Cícero, Pedro Paulo, Antônio José, Alex, Erik, Anísia e Kláudia. Em especial gostaria de agradecer aos professores Bruno, Daniel (desde a época do Mestrado) e Francisco Gilberto (desde a época do Mestrado) pela amizade durante esses anos de UFPI e por contribuírem para a realização deste momento, com palavras de apoio e incentivo.

Agradeço a todos os alunos do Doutorado. Em especial: Edimilson, Christopher, Alexandre, Atecio, Gustavo e Bruno que ingressaram comigo no ano de 2019.

Agradeço a todos os professores e funcionários do PGMAT e Departamento de Matemática. Em especial aos professores João Xavier e Barnabé pelas orientações, sugestões e incentivo.

”Eu me deitei e dormi; acordei, porque o
Senhor me sustentou.”

Salmos 3:5

Resumo

Neste trabalho, apresentamos dois métodos do ponto proximal para encontrar zeros de campos de vetores possivelmente não-monótonos que são escritos como diferença de dois campos monótonos maximais em variedades de Hadamard. O primeiro método é uma extensão natural do método do ponto proximal para diferença de funções convexas e que generaliza o método do ponto proximal para campos vetoriais monótonos maximais. O segundo método é uma versão inercial do primeiro e é novo inclusive no contexto monótono em variedades de Hadamard. Para ambos os métodos, provamos que, caso existam, os pontos de acumulação da sequência gerada pelos algoritmos são soluções do problema. Com hipóteses adicionais razoáveis, estabelecemos condições suficientes para limitação e convergência global da sequência. Alguns experimentos numéricos ilustram os resultados obtidos.

Palavras-chave: Método de ponto proximal; diferença de campos vetoriais monótonos maximais; variedades de Hadamard.

Abstract

In this work, we present two proximal point methods to find zeros of possibly non-monotone vector fields which are written as the difference of two maximal monotone fields in Hadamard manifolds. The first method is a natural extension of the proximal point method for difference of convex functions and which generalizes the proximal point method for maximal monotone vector fields. The second method is an inertial version of the first one and it is new even in the monotone context in Hadamard manifolds. For both methods, we prove that every cluster point of the sequences generated by the algorithms, if any, are solutions of the problem. Under reasonable additional hypotheses, we establish sufficient conditions for boundedness and global convergence of the sequence. Some numerical experiments illustrate the results obtained.

Keywords: Proximal point method; difference of maximal monotone vector fields; Hadamard manifolds.

Sumário

| | |
|--|-----------|
| Resumo | iv |
| Abstract | v |
| Introdução | 1 |
| 1 Preliminares | 7 |
| 1.1 Variedades Riemannianas | 7 |
| 1.2 Convexidade em Variedades Riemannianas | 13 |
| 1.3 Campos Vetoriais Monótonos em Variedades de Hadamard | 18 |
| 2 Método do ponto proximal | 26 |
| 2.1 Algoritmo MPP | 27 |
| 2.2 Análise de convergência | 30 |
| 2.2.1 Análise de convergência parcial | 33 |
| 2.2.2 Condições suficientes para limitação | 35 |
| 2.2.3 Condição suficiente para convergência total | 38 |
| 3 Método do ponto proximal inercial | 43 |
| 3.1 Algoritmo MPPI | 43 |
| 3.2 Análise de convergência: caso possivelmente não monótono | 45 |
| 3.2.1 Análise de convergência parcial | 49 |
| 3.2.2 Condições suficientes para a limitação | 50 |
| 3.2.3 Condição suficiente para convergência total | 53 |
| 3.3 Análise de convergência: caso monótono | 58 |
| 4 Experimentos numéricos | 62 |

| | |
|-----------------------------------|-----------|
| Sumário | vii |
| 5 Considerações finais | 74 |
| Referências Bibliográficas | 75 |

Introdução

Os operadores monótonos foram introduzidos pela primeira vez por (MINTY, 1960) no contexto de redes elétricas, e depois, na configuração de equações diferenciais parciais por (BROWDER, 1983). As principais ideias no estudo dos operadores monótonos ficaram claras na década de 70 pelos trabalhos de (BREZIS, 1973; BREZIS; HARAUX, 1976) e (ROCKAFELLAR, 1970, 1976). Nas décadas seguintes, operadores monótonos maximais rapidamente encontraram aplicações e agora desempenham um grande papel na otimização e análise convexa. Referimos a (BORWEIN, 2010) que faz um breve levantamento da história do assunto nos últimos 50 anos.

O problema de encontrar um zero de operadores monótonos maximais T , ou seja,

$$0 \in T(x) \tag{1}$$

inclui, como casos especiais, problemas de otimização e min-max, problemas de complementaridade e desigualdades variacionais. Ele encontra muitas aplicações importantes em campos científicos, como processamento de imagens, visão computacional, aprendizado de máquina e processamento de sinais. Por esta razão, nos últimos anos muita atenção tem sido dada ao desenvolvimento de métodos numéricos eficientes e implementáveis para resolver este problema em diferentes contextos; veja por exemplo (ROCKAFELLAR, 1976; ALVAREZ; ATTOUCH, 2001; CRUZ NETO et al., 2006; LI et al., 2009; WANG et al., 2015; ANSARI; BABU, 2020; ANSARI et al., 2020; ANSARI; BABU, 2021).

Uma das abordagens fundamentais para resolver (1) é o método do ponto proximal (MPP). Em 1976, o notável trabalho de (ROCKAFELLAR, 1976) considerou o método do ponto proximal, introduzido pela primeira vez por (MARTINET, 1970), para encontrar uma solução da inclusão (1). Em um espaço de Hilbert H , ele gera uma sequência $\{x^k\}$

que satisfaz

$$0 \in T(x^{k+1}) - \frac{1}{\mu_k}(x^k - x^{k+1}), \quad (2)$$

onde $\{\mu_k\}$ é uma sequência auxiliar de parâmetros positivos e x^0 é um ponto inicial arbitrário. Se o operador monótono maximal T é o subdiferencial de uma função convexa f , então o método do ponto proximal (2) torna-se

$$x^{k+1} = \arg \min_{x \in H} \left\{ f(x) + \frac{1}{2\mu_k} \|x - x^k\|^2 \right\}. \quad (3)$$

O MPP (3) encontra, sob hipóteses razoáveis, um ponto de mínimo de f . Em 2002, (FERREIRA; OLIVEIRA, 2002) propuseram uma extensão do método do ponto proximal (3) para o cenário não linear das variedades Riemannianas, mais precisamente, para uma variedade Riemanniana completa, simplesmente conexa de curvatura seccional não positiva (chamada variedade de Hadamard). Anos depois (LI et al., 2009) propuseram uma extensão do método do ponto proximal (2) na mesma configuração de variedades de Hadamard, provaram sob algumas hipóteses razoáveis que a sequência gerada por (2) converge para uma singularidade de um campo vetorial monótono. Desde então, muitos autores estudaram métodos de pontos proximais em variedades de Hadamard, veja por exemplo (ANSARI et al., 2019; BENTO et al., 2015; SOUZA, 2018; WANG et al., 2015, 2016) e em variedades Riemannianas mais gerais ou espaços não lineares; veja por exemplo (BAČÁK, 2013; BENTO et al., 2016; HUANG; WEI, 2022; LEUȘTEAN et al., 2018). Nos últimos anos, a otimização em variedades Riemannianas ganhou muito interesse tanto do ponto de vista teórico quanto de aplicação. Nesse sentido, o desenvolvimento de técnicas e algoritmos de otimização em variedades, especialmente para funções não diferenciáveis, tem aumentado amplamente. Do ponto de vista computacional, lidar com algoritmos em variedades Riemannianas tem uma vantagem em relação ao espaço Euclidiano porque problemas não convexos no sentido clássico podem se tornar convexos e problemas de otimização restrita podem ser vistos como irrestritos através da introdução de uma métrica Riemanniana apropriada tornando assim a versão Riemanniana de um método mais eficiente que sua versão Euclidiana. Por exemplo em (FERREIRA et al., 2019), problemas de otimização Euclidiano restrito com a função objetivo não convexa tendo um gradiente não Lipschitz, pode ser visto como um problema de otimização Riemanniana irrestrito com a função objetivo convexa tendo gradiente Lipschitz, com isso o

método do gradiente Riemanniano teve melhor desempenho computacional que o método do gradiente Euclidiano. O trabalho de (CRUZ NETO et al., 2006) apresenta alguns exemplos de operadores não monótonos na configuração Euclidiana que passaram a ser monótonos na configuração Riemanniana. Nas aplicações, destacamos o surgimento de imagens multivaloradas, dados que consistem em orientações em difração de elétrons retroespalhados e tensores de difusão em imagens de ressonância magnética; ver por exemplo (BERGMANN et al., 2021) e suas referências.

Em 2015, (SOUZA; OLIVEIRA, 2015) consideraram pela primeira vez, no contexto das variedades de Hadamard, uma análise de convergência do método do ponto proximal para uma importante classe de funções (possivelmente) não convexas que podem ser escritas como diferença de duas funções convexas (chamadas funções DC). No cenário Euclidiano, DCA (DC Algoritmos) foi o primeiro algoritmo para programação DC; veja (PHAM; SOUAD, 1986). A partir desse algoritmo, surgiram diversas variantes do DCA e diversos problemas teóricos e práticos foram descobertos ao longo dos anos, resultando em uma vasta literatura sobre o assunto; veja (THI; PHAM, 2018) para uma perspectiva histórica. Nas últimas décadas, a programação DC tornou-se muito popular e um grande número de artigos sobre o tema tem surgido; veja a recente revisão da literatura (OLIVEIRA, 2020). No entanto, esse assunto ainda não foi profundamente explorado no cenário Riemanniano.

Sabemos que a soma de dois operadores monótonos é um operador monótono, enquanto que a diferença de dois operadores monótonos não é necessariamente um operador monótono. Portanto, o problema de encontrar um zero da diferença de dois operadores monótonos pode ser um problema não monótono, dessa forma os métodos clássicos para operadores monótonos existentes na literatura não podem ser aplicados. O mesmo generaliza o problema de encontrar os pontos críticos da diferença de duas funções convexas e não foi estudado extensivamente mesmo na configuração linear; ver por exemplo, (ALIMOHAMMADY et al., 2014; ALIMOHAMMADY; RAMAZANNEJAD, 2016; MOUDAFI, 2008, 2015; NOOR et al., 2009).

Neste trabalho propomos dois métodos de ponto proximal para encontrar zeros de um campo vetorial (possivelmente) não monótono que pode ser escrito como a diferença de dois campos vetoriais monótonos maximais (chamados campos vetoriais componentes) em variedades de Hadamard, ou seja, propomos métodos do ponto proximal para o problema

de inclusão variacional

$$0 \in A(x) - B(x), \quad (4)$$

onde A e B são campos vetoriais monótonos maximais na variedade de Hadamard M , de dimensão finita. O Problema (4) é equivalente ao problema de encontrar $x \in M$ tal que

$$A(x) \cap B(x) \neq \emptyset. \quad (5)$$

O Problema (5) ainda não foi estudado no cenário de variedades de Hadamard e, uma vez que o subdiferencial de uma função convexa é um campo monótono maximal, claramente (5) generaliza o problema de encontrar pontos críticos de funções DC. Uma interessante aplicação da diferença de operadores monótonos em espaços lineares é dada por (ATTOUCH; THÉRA, 1996). Eles mencionam que a maioria das equações que surgem na física, economia, entre outras, podem ser escritas como segue

$$0 \in C(x) = A(x) + B(x), \quad (6)$$

onde A e B são possivelmente operadores multivalorados. A divisão do operador C na soma de dois operadores A e B geralmente tem um significado físico ou econômico, pois A e B podem ter propriedades muito distintas. Em vez de estudar diretamente (6), os autores em (ATTOUCH; THÉRA, 1996) propuseram uma transformação de dualidade equivalente e consideraram a seguinte diferença de operadores multivalorados

$$0 \in A^{-1}(y) - B^{-1}(-y)$$

oferecendo uma nova abordagem de dualidade para algumas questões centrais na teoria das desigualdades variacionais e operadores monótonos maximais.

Sabe-se que lidando com algoritmos DC, apenas resultados de convergência parcial são obtidos (no sentido de que cada ponto de acumulação da sequência gerada pelo método, se houver, é uma solução do problema DC) sob hipótese razoáveis. Além dos resultados de convergência parcial para o caso mais geral de diferença de campos vetoriais monótonos maximais, também apresentamos condições suficientes para limitação e convergência global do MPP que são novas mesmo para funções DC. Vale ressaltar que o MPP para diferença de campos vetoriais monótonos maximais generaliza os trabalhos (LI et al., 2009)

(tomando o campo vetorial da segunda componente como nulo) e (SOUZA; OLIVEIRA, 2015) (considerando os campos vetoriais monótonos maximais como subdiferenciais de funções convexas).

O método do ponto proximal calcula a cada iteração o conhecido operador resolvente introduzido por (MOREAU, 1965). Infelizmente, em muitos casos interessantes, a avaliação do operador resolvente é tão difícil quanto resolver o problema original (1). Por outro lado, em muitos problemas, o operador T pode ser escrito como a soma de dois operadores monótonos maximais, a saber, $T = A + B$, tal que o operador resolvente de cada componente é muito mais fácil de calcular do que o operador original T . Os chamados métodos *forward-backward* superam essa desvantagem combinando os resolventes de cada componente para encontrar uma solução do problema original; por exemplo o algoritmo de (DOUGLAS; RACHFORD, 1956) entre outros.

Já (POLYAK, 1964) introduziu o chamado método *heavy ball* para minimização de uma função convexa suave. A diferença em relação ao método gradiente é que a cada iteração, um ponto de extrapolação (que combina as iterações atuais e anteriores) é usado em vez da iteração atual. Essa pequena alteração melhora o desempenho do método gradiente. Dessa forma, (ALVAREZ; ATTOUCH, 2001) adaptaram essa ideia ao método do ponto proximal para operadores monótonos maximais que ficou conhecido como método do ponto proximal inercial (MMPI). No contexto Euclidiano, métodos inerciais foram considerados no caso particular de funções DC em (MAINGÉ; MOUDAFI, 2008; OLIVEIRA; TCHEOU, 2019). No entanto, até onde sabemos, não existe nenhum estudo do método do ponto proximal inercial para inclusão variacional envolvendo a diferença de dois campos vetoriais monótonos maximais em variedades de Hadamard.

Nesse sentido, o segundo método proposto é uma versão inercial do MPP para diferença de dois campos vetoriais monótonos maximais em variedades de Hadamard. Além disso, vale a pena mencionar que o MPPI ainda é novo em alguns casos particulares em variedades de Hadamard, a saber, se $B(x) = 0$ em (5) temos uma versão inercial-mMPPI do método do ponto proximal considerado por (LI et al., 2009) para encontrar uma singularidade de um campo vetorial monótono maximal; se $A(x) = \partial g(x)$ e $B(x) = \partial h(x)$ em (5), onde g e h são funções convexas, então nosso MPPI se torna uma versão inercial do método do ponto proximal para funções DC ($f(x) = g(x) - h(x)$) proposto por (SOUZA; OLIVEIRA, 2015); se $A(x) = \partial f(x)$ e $B(x) = 0$ em (4), onde f é uma função convexa,

então temos uma versão inercial do método do ponto proximal proposto por (FERREIRA; OLIVEIRA, 2002) para problemas de minimização convexa; e finalmente, se o parâmetro que aparece na direção do método inercial MPPI é zero, temos o MPP para diferença de campos vetoriais monótonos maximais.

Para o MPPI provamos que todo ponto de acumulação da sequência, se houver, é uma solução do problema de inclusão variacional. Além disso, apresentamos algumas condições suficientes para limitação e convergência total do MPPI que são novos mesmo para funções DC em variedades de Hadamard.

Apresentamos alguns experimentos numéricos para ilustrar os resultados obtidos. Os exemplos apresentados envolvem funções convexas e funções DC não convexas em uma variedade de Hadamard genuína (com curvatura não constante diferente de zero) e no cenário Euclidiano envolvendo diferença de operadores monótonos.

Este trabalho está organizado da seguinte forma. No Capítulo 1 relembramos alguns notações, definições e resultados preliminares sobre variedades Riemannianas, convexidade e campos vetoriais em variedades de Hadamard que serão usados ao longo deste trabalho. No Capítulo 2, apresentamos o MPP para a diferença de campos vetoriais monótonos maximais. No Capítulo 3, apresentamos o MPPI para a diferença de campos vetoriais monótonos maximais. No Capítulo 4, são apresentados experimentos numéricos para apoiar os resultados teóricos obtidos e ilustrar a viabilidade e eficiência dos métodos propostos. Finalmente, algumas conclusões são apresentadas.

Capítulo 1

Preliminares

Neste capítulo, apresentamos alguns conceitos, resultados e notações já existentes na literatura que darão suporte teórico em todo texto desta tese. Vamos apresentar algumas notações e propriedades fundamentais sobre geometria Riemanniana (veja (DO CARMO, 2015) e (SAKAI, 1996)), conceitos e resultados clássicos da teoria de otimização em variedades Riemannianas (veja (UDRISTE, 1994)), definições e notações de campos vetoriais em variedades de Hadamard (veja (LI et al., 2009) e (CRUZ NETO et al., 2000)), e as noções e resultados de resolventes de um campo vetorial em variedades de Hadamard (veja (LI et al., 2011)).

1.1 Variedades Riemannianas

Nesta seção, apresentamos algumas definições, propriedades e notações fundamentais de variedades Riemannianas, que podem ser encontradas em (DO CARMO, 2015), (SAKAI, 1996) e (UDRISTE, 1994). Mais precisamente, veja esses conceitos e resultados em variedades de Hadamard em (DO CARMO, 2015, Capítulo VII) e (SAKAI, 1996, Capítulo V).

Seja M uma variedade diferenciável m -dimensional e conexa. Assumimos que M pode ser dotada de uma métrica Riemanniana $\langle \cdot, \cdot \rangle$, com a norma correspondente indicada por $\| \cdot \|$, para tornar-se uma variedade Riemanniana.

Definição 1. *O espaço tangente de M em $p \in M$ é denotado por $T_p M$ e o fibrado tangente de M é dado por $TM = \cup_{p \in M} T_p M$, que é naturalmente uma variedade. Um campo vetorial X em M é uma aplicação $X : M \rightarrow TM$ que associa a cada $p \in M$*

um $X(\mathbf{p}) \in T_{\mathbf{p}}M$. O campo vetorial X é diferenciável se a aplicação $X : M \rightarrow TM$ é diferenciável. O conjunto de todos os campos vetoriais em M de classe C^∞ é denotado por $\mathcal{X}(M)$.

Definição 2. Seja $f : M \rightarrow \mathbb{R}$ uma função de classe C^1 . Dado $\mathbf{x} \in M$, chamamos de gradiente de f em \mathbf{x} , o único campo vetorial $\text{grad } f(\mathbf{x})$ tal que

$$\langle \text{grad } f(\mathbf{x}), \mathbf{v} \rangle = \nabla_{\mathbf{v}} f = df_{\mathbf{x}}(\mathbf{v}), \quad \forall \mathbf{v} \in T_{\mathbf{x}}M.$$

Definição 3. Dada uma função $f : M \rightarrow \mathbb{R}$ de classe C^2 sua Hessiana em $\mathbf{x} \in M$ é a aplicação linear $\text{Hess } f(\mathbf{x}) : T_{\mathbf{x}}M \rightarrow T_{\mathbf{x}}M$ dada por

$$\text{Hess } f(\mathbf{x})\mathbf{v} = \nabla_{\mathbf{v}} \text{grad } f = d(\text{grad } f)_{\mathbf{x}}(\mathbf{v}), \quad \forall \mathbf{v} \in T_{\mathbf{x}}M.$$

Definição 4. Seja $\gamma : [a, b] \rightarrow M$ uma curva suave por partes ligando \mathbf{p} a \mathbf{q} pontos de M , isto é, $\gamma(a) = \mathbf{p}$ e $\gamma(b) = \mathbf{q}$. O comprimento da curva γ usando a métrica, é dado por $L(\gamma) = \int_a^b \|\gamma'(t)\| dt$. Assim, a distância Riemanniana $d(\mathbf{p}, \mathbf{q})$, é definida minimizando este comprimento sobre o conjunto de todas essas curvas ligando os pontos \mathbf{p} e \mathbf{q} . A função distância Riemanniana $d : M \times M \rightarrow \mathbb{R}$ induz a topologia original em M .

Definição 5. Seja ∇ a conexão Levi-Civita associada a $(M, \langle \cdot, \cdot \rangle)$. Seja $\gamma : I \rightarrow M$ uma curva suave em M . Diz-se que um campo vetorial X é paralelo ao longo de γ se $\nabla_{\gamma'} X = 0$. Se γ' em si é paralelo ao longo de γ , então dizemos que γ é uma geodésica e, neste caso, $\|\gamma'\|$ é constante. Quando $\|\gamma'\| = 1$, γ é chamada normalizada.

Observação 1. Dado que a equação geodésica $\nabla_{\gamma'} \gamma' = 0$ é uma equação diferencial ordinária não linear de segunda ordem, então a geodésica $\gamma(\cdot) = \gamma_{\mathbf{v}}(\cdot, \mathbf{p})$ é determinada por sua posição $\mathbf{p} \in M$ e sua velocidade $\mathbf{v} \in T_{\mathbf{p}}M$ em \mathbf{p} .

Definição 6. A restrição de uma geodésica $\gamma : I \rightarrow M$ a um intervalo fechado e limitado $[a, b] \subset I$ é chamado de segmento geodésico e denotaremos o conjunto de todos os segmentos geodésicos ligando dois pontos \mathbf{p} e \mathbf{q} por $\Gamma_{\mathbf{p}\mathbf{q}}$. Diz-se que um segmento geodésico ligando \mathbf{p} a \mathbf{q} em M é minimal se seu comprimento for igual a $d(\mathbf{p}, \mathbf{q})$. Nesse caso, diz-se que a geodésica é minimizante.

Definição 7. Uma variedade Riemanniana M é (geodesicamente) completa em $\mathbf{p} \in M$ se toda geodésica $\gamma(t)$ que parte de \mathbf{p} está definida para todo $t \in \mathbb{R}$. A variedade Riemanniana

M é chamada (geodesicamente) completa se é (geodesicamente) completa em todo ponto $p \in M$.

Observação 2. *Seja M uma variedade Riemanniana completa. Prova-se que para todo $a \in \mathbb{R}$, com $a > 0$ a igualdade $\gamma_{av}(t) = \gamma_v(at)$ é válida para todo $t \in \mathbb{R}$. De fato, basta usar o Lema 2.6 em (DO CARMO, 2015) e o fato que M é completa.*

Definição 8. *Seja M uma variedade Riemanniana completa e $p \in M$, a aplicação exponencial $\exp_p : T_p M \rightarrow M$ é definida por*

$$\exp_p v = \gamma_v(1, p),$$

onde $\gamma(\cdot) = \gamma_v(\cdot, p)$ é a geodésica partindo de p com vetor velocidade v .

Observação 3. *Segue da **Observação 2** que $\exp_p tv = \gamma_v(t, p)$. Note que a aplicação \exp_p é diferenciável em $T_p M$ para todo $p \in M$.*

O resultado que veremos a seguir conhecido como **Teorema de Hopf-Rinow** (veja por exemplo (DO CARMO, 2015) e (SAKAI, 1996)) assegura que qualquer dois pontos p e q em M podem ser ligados por um (não necessariamente único) segmento geodésico minimal. Além disso, $(M; d)$ é um espaço métrico completo e, subconjuntos fechados e limitados são compactos.

Teorema 1. (Teorema de Hopf-Rinow) *Seja M uma variedade Riemanniana. As seguintes condições são equivalentes:*

- (i) *Para todo ponto $p \in M$, \exp_p está definida em todo o $T_p M$.*
- (ii) *Os subconjuntos limitados e fechados de M são compactos.*
- (iii) *(M, d) é um espaço métrico completo, onde d é a distância Riemanniana.*
- (iv) *M é (geodesicamente) completa em $p \in M$.*
- (v) *M é (geodesicamente) completa.*

Além disso, cada um dos itens acima (i) – (v) implicam na seguinte condição:

- (vi) *Para quaisquer dois pontos $p, q \in M$ existe um segmento geodésico γ (minimal) ligando p a q tal que $L(\gamma) = d(p, q)$, isto é, γ é uma geodésica minimal.*

Demonstração. Veja (DO CARMO, 2015, Teorema 2.8) e (SAKAI, 1996, Teorema 1.1). \square

Seja $\gamma : [\mathbf{a}, \mathbf{b}] \rightarrow M$ um segmento geodésico ligando $\mathbf{p} := \gamma(\mathbf{a})$ a $\mathbf{q} = \gamma(\mathbf{b})$ em M . Para cada $\mathbf{t} \in [\mathbf{a}, \mathbf{b}]$, a conexão Levi-Civita ∇ induz uma isometria linear, relativa a $\langle \cdot, \cdot \rangle$: $P_{\gamma, \gamma(\mathbf{b}), \gamma(\mathbf{a})} : T_{\gamma(\mathbf{a})}M \rightarrow T_{\gamma(\mathbf{b})}M$, que é definida por

$$P_{\gamma, \gamma(\mathbf{b}), \gamma(\mathbf{a})}(\mathbf{v}) = \mathbf{V}(\gamma(\mathbf{b})), \quad \mathbf{a}, \mathbf{b} \in \mathbb{R}, \mathbf{v} \in T_{\gamma(\mathbf{a})}M,$$

onde \mathbf{V} é o único campo vetorial satisfazendo $\nabla_{\gamma'(t)}^{\mathbf{V}} = 0$ para todo t e $\mathbf{V}(\gamma(\mathbf{a})) = \mathbf{v}$.

Definição 9. A aplicação $P_{\gamma, \gamma(\mathbf{b}), \gamma(\mathbf{a})}$ é chamada o transporte paralelo ao longo de γ de $\gamma(\mathbf{a}) = \mathbf{p}$ a $\gamma(\mathbf{b}) = \mathbf{q}$.

Observação 4. Para quaisquer $\mathbf{a}, \mathbf{b}, \mathbf{b}_1, \mathbf{b}_2 \in \mathbb{R}$ valem as seguintes condições:

(i) $P_{\gamma, \gamma(\mathbf{b}_2), \gamma(\mathbf{b}_1)} \circ P_{\gamma, \gamma(\mathbf{b}_1), \gamma(\mathbf{a})} = P_{\gamma, \gamma(\mathbf{b}_2), \gamma(\mathbf{a})}$;

(ii) a aplicação inversa de $P_{\gamma, \gamma(\mathbf{b}), \gamma(\mathbf{a})}$ é denotada por $P_{\gamma, \gamma(\mathbf{b}), \gamma(\mathbf{a})}^{-1} : T_{\gamma(\mathbf{b})}M \rightarrow T_{\gamma(\mathbf{a})}M$, com

$$P_{\gamma, \gamma(\mathbf{b}), \gamma(\mathbf{a})}^{-1} = P_{\gamma, \gamma(\mathbf{a}), \gamma(\mathbf{b})};$$

(iii) no caso em que γ é uma geodésica minimal ligando os pontos \mathbf{p} e \mathbf{q} em M o transporte paralelo ao longo de γ de \mathbf{p} a \mathbf{q} será denotado por $P_{\mathbf{q}, \mathbf{p}} : T_{\mathbf{p}}M \rightarrow T_{\mathbf{q}}M$.

Definição 10. Sejam $X, Y, Z \in \mathcal{X}(M)$, o tensor curvatura R é definido por

$$R(X, Y) = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[Y, X]} Z,$$

onde $[X, Y] = YX - XY$ (operação colchete, veja Capítulo 0 em (DO CARMO, 2015)).

Definição 11. Sejam $X, Y \in \mathcal{X}(M)$, a curvatura seccional com respeito a X e Y é dada por

$$K(X, Y) = (\langle R(X, Y)Y, X \rangle) / (\|X\|^2 \|Y\|^2 - \langle X, Y \rangle^2),$$

onde $\|X\|^2 = \langle X, X \rangle$.

Definição 12. Uma variedade Riemanniana completa e simplesmente conexa de curvatura seccional não positiva é chamada de variedade de Hadamard.

Vejamos agora alguns exemplos de variedades de Hadamard, ou seja, variedades Riemannianas de curvaturas seccional K não positiva.

Exemplo 1. Se $M = \mathbb{R}^m$ é o espaço Euclidiano m -dimensional com a métrica usual, então a curvatura de M é constante $K = 0$.

Exemplo 2. Se $M = \mathbb{H}^m$ é o espaço Hiperbólico com a métrica $g_{ij} = \frac{\delta_{ij}}{x_m^2}$, então a curvatura de M é constante $K = -1$.

Exemplo 3. Sejam \mathbb{P}^n o conjunto das matrizes simétricas, \mathbb{P}_+^n o cone das matrizes simétricas semidefinidas positivas e \mathbb{P}_{++}^n o cone das matrizes simétricas definidas positivas $n \times n$. Seja $M := (\mathbb{P}_{++}^n, \langle \cdot, \cdot \rangle)$ a variedade Riemanniana dotada da métrica Riemanniana dada por

$$\langle U, V \rangle_X := \text{tr}(VX^{-1}UX^{-1}), \quad X \in M, \quad U, V \in T_X M,$$

onde $\text{tr}(X)$ denota o traço de $X \in \mathbb{P}^n$ e $T_X M \approx \mathbb{P}^n$, com a correspondente norma denotada por $\| \cdot \|$; veja (ROTHAUS, 1960). Neste caso, para quaisquer $X, Y \in M$ a única geodésica ligando esses dois pontos é dada por:

$$\gamma(t) = X^{1/2} (X^{-1/2} Y X^{-1/2})^t X^{1/2}, \quad \gamma'(0) = X^{1/2} \ln(X^{-1/2} Y X^{-1/2}) X^{1/2},$$

$t \in [0, 1]$, veja, por exemplo, (NESTEROV; TODD, 2002) e (BENTO et al., 2015). Mais precisamente, M é uma variedade de Hadamard, veja, por exemplo, (LANG, 2012, Teorema 1.2., página 325). Pode-se calcular a curvatura de M e verificar que ela possui curvatura não constante; veja (LENGLET et al., 2006, página 428).

A aplicação exponencial $\exp_X : T_X M \rightarrow M$ e $\exp_X^{-1} : M \rightarrow T_X M$ sua inversa são dadas da seguinte forma

$$\exp_X V = X^{1/2} e^{X^{-1/2} V X^{-1/2}} X^{1/2}, \quad \exp_X^{-1} Y = X^{1/2} \ln(X^{-1/2} Y X^{-1/2}) X^{1/2},$$

veja (BHATIA, 2009, Capítulo 6) e (BOUMAL, 2022, Capítulo 11).

O teorema a seguir garante que uma variedade de Hadamard tem a mesma topologia e estrutura diferenciável dos espaços Euclidianos \mathbb{R}^m .

Teorema 2. (Teorema de Cartan-Hadamard) Sejam M uma variedade de Hadamard, então M é difeomorfa ao espaço Euclidiano \mathbb{R}^m , $m = \dim M$. Mais precisamente $\exp_p : T_p M \rightarrow M$, é um difeomorfismo de classe C^∞ para cada $p \in M$.

Demonstração. Veja (DO CARMO, 2015, Teorema 3.1) e (SAKAI, 1996, Teorema 4.1 página 221). \square

Observação 5. *Sejam M uma variedade de Hadamard e $p \in M$. Pelo Teorema de Cartan-Hadamard podemos definir a inversa da aplicação exponencial $\exp_p^{-1} : M \rightarrow T_p M$ e obtém-se a seguinte relação entre a distância Riemanniana e a inversa da aplicação exponencial:*

$$d(p, q) = \|\exp_p^{-1} q\|.$$

Como \exp_p^{-1} é uma aplicação de classe C^∞ , segue da relação acima que $d^2(p, \cdot)$ é também de classe C^∞ . Foi provado em (SAKAI, 1996, Proposição 4.8 página 108) que

$$\text{grad} \frac{1}{2} d^2(p, q) = -\exp_p^{-1} q.$$

Definição 13. *Um triângulo geodésico $\Delta(p_1, p_2, p_3)$ de uma variedade Riemanniana é um conjunto que consiste em três pontos distintos p_1, p_2 e p_3 chamados vértices do triângulo e três segmentos de geodésicas minimizantes chamadas lados do triângulo $\gamma_i : [0, l_i] \rightarrow M$ ligando esses três pontos com $i = 1, 2, 3$. O ângulo $\alpha_i := \angle(\gamma_i'(0), -\gamma_{i-1}'(l_{i-1}))$ é chamado ângulo interno do vértice correspondente.*

Além de propriedades topológicas e diferenciáveis, algumas propriedades geométricas similares às dos espaços Euclidianos também são obtidas em variedades de Hadamard, como o teorema a seguir.

Teorema 3. (Teorema de Comparação para Triângulos)

Seja $\Delta(p_1, p_2, p_3)$ um triângulo geodésico, $\gamma_i : [0, l_i] \rightarrow M$ a geodésica ligando p_i a p_{i+1} , $l_i = L(\gamma_i)$ e $\alpha_i := \angle(\gamma_i'(0), -\gamma_{i-1}'(l_{i-1}))$ para cada $i = 1, 2, 3$. Então,

$$\alpha_1 + \alpha_2 + \alpha_3 \leq \pi,$$

$$l_i^2 + l_{i+1}^2 - 2l_i l_{i+1} \cos \alpha_{i-1} \leq l_{i-1}^2. \quad (1.1)$$

Em termos da distância e da aplicação exponencial, a Desigualdade (1.1) pode ser reescrita da seguinte forma:

$$d^2(p_i, p_{i+1}) + d^2(p_{i+1}, p_{i+2}) - 2\langle \exp_{p_{i+1}}^{-1} p_i, \exp_{p_{i+1}}^{-1} p_{i+2} \rangle \leq d^2(p_{i-1}, p_i), \quad (1.2)$$

desde que $\langle \exp_{\mathbf{p}_{i+1}}^{-1} \mathbf{p}_i, \exp_{\mathbf{p}_{i+1}}^{-1} \mathbf{p}_{i+2} \rangle = d(\mathbf{p}_{i+1}, \mathbf{p}_i) d(\mathbf{p}_{i+1}, \mathbf{p}_{i+2}) \cos \alpha_{i+1}$.

Demonstração. Veja (SAKAI, 1996, Proposição 4.5 página 223). \square

Usando as propriedades do transporte paralelo e da aplicação exponencial, obtemos a seguinte proposição que será usada com frequência nas próximas seções.

Proposição 1. *Seja M uma variedade de Hadamard. Sejam $\mathbf{x}, \mathbf{y}, \bar{\mathbf{x}} \in M$ e $\{\mathbf{x}^k\}, \{\mathbf{y}^k\} \subset M$ tais que $\mathbf{x}^k \rightarrow \mathbf{x}$ e $\mathbf{y}^k \rightarrow \mathbf{y}$. Então, as seguintes condições são válidas:*

(i) *Para todo $\mathbf{y} \in M$, temos $\exp_{\mathbf{x}^k}^{-1} \mathbf{y} \rightarrow \exp_{\mathbf{x}}^{-1} \mathbf{y}$ e $\exp_{\mathbf{y}^k}^{-1} \mathbf{x}^k \rightarrow \exp_{\mathbf{y}}^{-1} \mathbf{x}$.*

(ii) *Se $\mathbf{v}^k \in T_{\mathbf{x}^k}M$ e $\mathbf{v}^k \rightarrow \mathbf{v}$, então $\mathbf{v} \in T_{\mathbf{x}}M$.*

(iii) *Dados, $\mathbf{u}^k, \mathbf{v}^k \in T_{\mathbf{x}^k}M$ e $\mathbf{u}, \mathbf{v} \in T_{\mathbf{x}}M$, se $\mathbf{u}^k \rightarrow \mathbf{u}$ e $\mathbf{v}^k \rightarrow \mathbf{v}$, então*

$$\langle \mathbf{u}^k, \mathbf{v}^k \rangle \rightarrow \langle \mathbf{u}, \mathbf{v} \rangle.$$

(iv) *Para todo $\mathbf{u} \in T_{\bar{\mathbf{x}}}M$, a função $F : M \rightarrow TM$ definida por $F(\mathbf{x}) = P_{\mathbf{x}, \bar{\mathbf{x}}} \mathbf{u}$ para cada $\mathbf{x} \in M$ é contínua em M .*

(v) *$\exp_{\mathbf{y}^k}^{-1} \mathbf{x}^k \rightarrow \exp_{\mathbf{y}}^{-1} \mathbf{x}$.*

Demonstração. Veja (LI et al., 2009) e (BATISTA et al., 2020). \square

1.2 Convexidade em Variedades Riemannianas

Nesta seção, apresentamos algumas definições, notações e propriedades de convexidade em variedades Riemannianas. Os resultados apresentados, bem como outros resultados de convexidade em variedades Riemannianas podem ser encontrados, por exemplo, em (BOUMAL, 2022), (FERREIRA; OLIVEIRA, 2002), (UDRISTE, 1994) e (CRUZ NETO et al., 1998). Muitos dos resultados a seguir são válidos para variedades Riemannianas, mas focaremos apenas nas variedades de Hadamard que será o ambiente a ser desenvolvido o trabalho de tese.

Definição 14. *Um subconjunto $C \subset M$ é dito convexo se, para quaisquer pontos $\mathbf{p}, \mathbf{q} \in C$, a geodésica ligando o ponto \mathbf{p} a \mathbf{q} está contida em C , isto é, se $\gamma : [a, b] \rightarrow M$ é uma geodésica tal que $\gamma(\mathbf{a}) = \mathbf{p}$ e $\gamma(\mathbf{b}) = \mathbf{q}$, então $\gamma((1-t)\mathbf{a} + t\mathbf{b}) \in C$ para todo $t \in [0, 1]$, isto é, $\gamma([a, b]) \subset C$.*

Definição 15. Uma função $f : M \rightarrow \mathbb{R}$ é convexa se, para todo segmento de geodésica $\gamma : [a, b] \rightarrow M$ ligando p a q em M , a composição $f \circ \gamma : [a, b] \rightarrow \mathbb{R}$ é convexa, isto é,

$$f(\gamma(t)) = (f \circ \gamma)((1-t)s_1 + ts_2) \leq (1-t)(f \circ \gamma)(s_1) + (f \circ \gamma)(s_2),$$

para todo $t \in [0, 1]$ e quaisquer $s_1, s_2 \in [a, b]$, tais que $\gamma(a) = p$ e $\gamma(b) = q$. Se a desigualdade acima for estrita para $s_1 \neq s_2$, então f é estritamente convexa.

Definição 16. A função $f : M \rightarrow \mathbb{R}$ é fortemente convexa se existir uma constante positiva m , para todo segmento de geodésica $\gamma : [a, b] \rightarrow M$ ligando p a q em M , a composição $f \circ \gamma : [a, b] \rightarrow \mathbb{R}$ é fortemente convexa, isto é,

$$f(\gamma(t)) = (f \circ \gamma)((1-t)s_1 + ts_2) \leq (1-t)(f \circ \gamma)(s_1) + (f \circ \gamma)(s_2) - \frac{m}{2} \|\gamma'(t)\| (1-t)t$$

para todo $t \in [0, 1]$ e quaisquer $s_1, s_2 \in [a, b]$, tais que $\gamma(a) = p$ e $\gamma(b) = q$.

Condição de primeira ordem para convexidade de uma função $f : M \rightarrow \mathbb{R}$ de classe C^1 .

Teorema 4. A função $f : M \rightarrow \mathbb{R}$ de classe C^1 é convexa se, e somente se, para qualquer ponto $x \in M$ e toda geodésica $\gamma : [0, +\infty) \rightarrow M$, com $\gamma(0) = x$, vale a desigualdade

$$f(\gamma(t)) - f(x) \geq t \langle \text{grad } f(x), \gamma'(0) \rangle.$$

Demonstração. Veja (CRUZ NETO et al., 1998, Teorema 3.1). □

Condição de segunda ordem para convexidade de uma função $f : M \rightarrow \mathbb{R}$ de classe C^2 .

Teorema 5. Uma função $f : M \rightarrow \mathbb{R}$ de classe C^2 é convexa se, e somente se, para todo $x \in M$, a Hessiana de f é semi-definida positiva, isto é,

$$\langle \text{Hess } f(x)v, v \rangle \geq 0, \quad \forall v \in T_x M.$$

Demonstração. Veja (CRUZ NETO et al., 1998, Teorema 3.2) e (UDRISTE, 1994, Teorema 6.2 página 82). □

Exemplo 4. Seja $M := (\mathbb{P}_{++}^n, \langle \cdot, \cdot \rangle)$ a variedade de Hadamard do Exemplo 3, onde \mathbb{P}_{++}^n é o cone das matrizes simétricas definidas positivas $n \times n$ e a métrica Riemanniana é dada por

$$\langle U, V \rangle_X := \text{tr}(VX^{-1}UX^{-1}), \quad X \in M, \quad U, V \in T_X M.$$

Para quaisquer $X, Y \in M$ a única geodésica ligando esses dois pontos é dada por:

$$\gamma(t) = X^{1/2} (X^{-1/2} Y X^{-1/2})^t X^{1/2}, \quad \gamma'(0) = X^{1/2} \ln(X^{-1/2} Y X^{-1/2}) X^{1/2},$$

$t \in [0, 1]$, cuja aplicação exponencial e sua inversa são dadas da seguinte forma

$$\exp_X V = X^{1/2} e^{X^{-1/2} V X^{-1/2}} X^{1/2}, \quad \exp_X^{-1} Y = X^{1/2} \ln(X^{-1/2} Y X^{-1/2}) X^{1/2}.$$

O gradiente e a Hessiana de uma função duas vezes diferenciável $f : \mathbb{P}_{++}^n \rightarrow \mathbb{R}$ são dados por

$$\text{grad } f(X) = Xf'(X)X \text{ e } \text{Hess } f(X)V = Xf''(X)VX + \frac{1}{2}(Vf'(X)X + Xf'(X)V)$$

onde $V \in T_X M$, $f'(X)$ e $f''(X)$ são o gradiente e a Hessiana Euclidiana de f respectivamente; veja por exemplo (ALMEIDA et al., 2020) e (BENTO et al., 2015).

Veremos agora alguns exemplos de funções convexas e não convexas em \mathbb{P}_{++}^n .

(i) $f_1(X) = a \ln(\det X)^2$, $a \in \mathbb{R}_{++}$ é convexa. De fato, o gradiente e a Hessiana euclidiana de f_1 são dados por:

$$f_1'(X) = 2a \ln(\det X)X^{-1} \text{ e } f_1''(X)V = 2a \text{tr}(X^{-1}V)X^{-1} - 2a \ln(\det X)X^{-1}VX^{-1}.$$

Então, $\text{grad } f_1(X) = 2a \ln(\det X)X$ e $\text{Hess } f_1(X)V = 2a \text{tr}(X^{-1}V)X$.

Assim, para cada X e todo $V \in T_X M$, temos

$$\langle \text{Hess } f_1(X)V, V \rangle = 2a [\text{tr}(X^{-1}V)]^2 \geq 0.$$

Portanto, f_1 é uma função convexa pelo Teorema 5.

(ii) $f_2(X) = b \ln(\det X)^4$, $b \in \mathbb{R}_{++}$ é convexa. De fato, o gradiente e a Hessiana Euclidiana de f_2 são dados por: $f_2'(X) = 4b \ln(\det X)^3 X^{-1}$ e

$$f_2''(X)V = 12b \ln(\det X)^2 \operatorname{tr}(X^{-1}V)X^{-1} - 4b \ln(\det X)^3 X^{-1}VX^{-1}.$$

Então, $\operatorname{grad} f_2(X) = 4b \ln(\det X)^3 X$ e $\operatorname{Hess} f_2(X)V = 12b \ln(\det X)^2 \operatorname{tr}(X^{-1}V)X$.

Assim, para cada X e todo $V \in T_X M$, temos

$$\langle \operatorname{Hess} f_2(X)V, V \rangle = 12b \ln(\det X)^2 [\operatorname{tr}(X^{-1}V)]^2 \geq 0.$$

Portanto, f_2 é uma função convexa pelo Teorema 5.

(iii) $f_3(X) = b \ln(\det X)^4 - a \ln(\det X)^2$, $a, b \in \mathbb{R}_{++}$ não é convexa. De fato, segue de (i) e (ii) que

$$\operatorname{grad} f_3(X) = 2 \ln(\det X) \left(2b \ln(\det X)^2 - a \right) X$$

e

$$\operatorname{Hess} f_3(X)V = 2 \operatorname{tr}(X^{-1}V) \left(6b \ln(\det X)^2 - a \right) X.$$

Assim, para cada $X \in \mathbb{P}_{++}^n$ e todo $V \in T_X M$, temos

$$\langle \operatorname{Hess} f_3(X)V, V \rangle = 2 [\operatorname{tr}(X^{-1}V)]^2 \left(6b \ln(\det X)^2 - a \right) < 0,$$

desde que $\det X \in [e^{-\sqrt{\frac{a}{6b}}}, e^{\sqrt{\frac{a}{6b}}}]$. Logo, f_3 não é convexa pelo Teorema 5.

(iv) $f_4(X) = b \ln(\det X)$, $b \in \mathbb{R}_{++}$ é convexa. De fato, o gradiente e a Hessiana Euclidiana de f_4 são dados por: $f_4'(X) = bX^{-1}$ e $f_4''(X)V = -bX^{-1}VX^{-1}$. Segue que,

$$\operatorname{grad} f_4(X) = bX \text{ e } \operatorname{Hess} f_4(X)V = 0.$$

Então, para cada $X \in \mathbb{P}_{++}^n$ e todo $V \in T_X M$, temos que $\langle \operatorname{Hess} f_4(X)V, V \rangle = 0$.

Portanto, f_4 é uma função convexa pelo Teorema 5.

(v) $f_5(X) = e^{a \operatorname{tr}(X)}$, $a \in \mathbb{R}_{++}$ é convexa. De fato, o gradiente e a Hessiana euclidiana de f_5 são dados por: $f_5'(X) = ae^{a \operatorname{tr}(X)}$ e $f_5''(X)V = a^2 e^{a \operatorname{tr}(X)} V$. Então,

$$\operatorname{grad} f_5(X) = ae^{a \operatorname{tr}(X)} X^2 \text{ e } \operatorname{Hess} f_5(X)V = a^2 e^{a \operatorname{tr}(X)} XVX + \frac{1}{2} ae^{a \operatorname{tr}(X)} (VX + XV).$$

Assim, para cada $X \in \mathbb{P}_{++}^n$ e todo $V \in T_X M$, temos que

$$\langle \text{Hess } f_5(X)V, V \rangle = a e^{a \text{tr}(X)} \text{tr}((aI + X^{-1})V^2) \geq 0.$$

Portanto, f_5 é uma função convexa pelo Teorema 5.

(vi) $f_6(X) = e^{\text{tr}(X)} - \ln(\det X)$ não é convexa. De fato, segue de (iv) e (v) que

$$\text{grad } f_6(X) = e^{\text{tr}(X)} X^2 - X$$

e

$$\text{Hess } f_6(X)V = 2^{\text{tr}X} X V X + \frac{1}{2} e^{\text{tr}X} (V X + X V) - 2V.$$

Para cada $X \in \mathbb{P}_{++}^n$ e todo $V \in T_X M$, temos que

$$\langle \text{Hess } f_6(X)V, V \rangle = e^{\text{tr}(X)} \text{tr}((I + X^{-1})V^2) - 2\text{tr}((VX^{-1})^2).$$

Definição 17. Seja $f : M \rightarrow \mathbb{R}$ uma função convexa. O subdiferencial de f em $x \in M$ é definido por

$$\partial f(x) = \{u \in T_x M; \langle u, \exp_x^{-1} y \rangle \leq f(y) - f(x), \forall y \in M\}.$$

O vetor $u \in \partial f(x)$ é chamado o subgradiente de f em x . Em particular se f é diferenciável, temos que $\partial f(x) = \{\text{grad } f(x)\}$, para todo $x \in M$; veja (BENTO et al., 2017).

Proposição 2. Seja $\{x^k\}$ uma sequência limitada em M . Se a sequência $\{v^k\}$ é tal que $v^k \in \partial f(x^k)$ para cada $k \in \mathbb{N}$, então $\{v^k\}$ é também limitada.

Demonstração. Veja (BENTO; MELO, 2012, Proposição 3.2). □

A proposição a seguir mostra que o subdiferencial $\partial f(x)$ de f em $x \in M$ não é vazio.

Proposição 3. Seja M uma variedade de Hadamard e seja $f : M \rightarrow \mathbb{R}$ uma função convexa. Então, para todo $x \in M$, existe $s \in T_x M$ tal que

$$f(y) \geq f(x) + \langle s, \exp_x^{-1} y \rangle, \forall y \in M.$$

Demonstração. Veja (FERREIRA; OLIVEIRA, 2002, Teorema 3.3). □

Seja $f : M \rightarrow \mathbb{R}$ uma função convexa. Dado $\lambda > 0$, a regularização de Moreau-Yosida de f é a função $f_\lambda : M \rightarrow \mathbb{R}$ dada por

$$f_\lambda(x) := \inf_{y \in M} \{f(y) + \frac{\lambda}{2} d^2(x, y)\}.$$

Proposição 4. *Seja $f : M \rightarrow \mathbb{R}$ uma função convexa, então para todo $x \in M$ e $\lambda > 0$, existe um único ponto, denotado por $p_\lambda(x)$, tal que*

$$f(p_\lambda(x)) + \frac{\lambda}{2} d^2(p_\lambda(x), x) = f_\lambda(x)$$

caracterizada por $\lambda(\exp_{p_\lambda(x)}^{-1} x) \in \partial f(p_\lambda(x))$.

Demonstração. Veja (FERREIRA; OLIVEIRA, 2002, Lema 4.2). □

1.3 Campos Vetoriais Monótonos em Variedades de Hadamard

Nesta seção, apresentamos algumas definições, notações e propriedades de campos vetoriais em variedades de Hadamard; veja por exemplo (CRUZ NETO et al., 2000), (LI et al., 2009) e (LI et al., 2011) para mais detalhes.

Seja $\mathcal{X}(M)$ o conjunto de todos os campos vetoriais $A : M \rightarrow 2^{T^M}$ de valores múltiplos, de modo que $A(x) \subset T_x M$ para cada $x \in M$ e o domínio $D(A)$ de A é fechado e convexo, onde o domínio $D(A)$ de A é definido por

$$D(A) = \{x \in M; A(x) \neq \emptyset\}.$$

Na definição abaixo, o conceito de campos vetoriais (estrita e fortemente) monótono em variedades Riemannianas foi introduzido pela primeira vez em (NÉMETH, 1999) para o caso de valor único e em (CRUZ NETO et al., 2000) para o caso de valores múltiplos.

Definição 18. (LI et al., 2009) *Seja $A \in \mathcal{X}(M)$. Então, A é dito*

(i) *monótono se a seguinte condição é válida para quaisquer $x, y \in D(A)$:*

$$\langle u, \exp_x^{-1} y \rangle \leq \langle v, -\exp_y^{-1} x \rangle, \quad \forall u \in A(x) \text{ e } \forall v \in A(y); \quad (1.3)$$

(ii) *estritamente monótono se (1.3) é válida com a desigualdade estrita para quaisquer $x, y \in D(A)$ com $x \neq y$, isto é,*

$$\langle u, \exp_x^{-1} y \rangle < \langle v, -\exp_y^{-1} x \rangle, \quad \forall u \in A(x) \text{ e } \forall v \in A(y); \quad (1.4)$$

(iii) *fortemente monótono se existe $\rho > 0$ tal que, para quaisquer $x, y \in D(A)$, temos*

$$\langle u, \exp_x^{-1} y \rangle - \langle v, -\exp_y^{-1} x \rangle \leq -\rho d^2(x, y), \quad \forall u \in A(x) \text{ e } \forall v \in A(y); \quad (1.5)$$

(iv) *monótono maximal se é monótono e a seguinte implicação é válida para quaisquer $x \in D(A)$ e $u \in T_x M$:*

$$\langle u, \exp_x^{-1} y \rangle \leq \langle v, -\exp_y^{-1} x \rangle, \quad \forall y \in D(A) \text{ e } v \in A(y) \implies u \in A(x). \quad (1.6)$$

Vejamos agora alguns exemplos de campos vetoriais monótonos e não monótonos.

Exemplo 5. (CRUZ NETO et al., 2000, Proposição 4.2)

Se $f : M \rightarrow \mathbb{R}$ é convexa, então ∂f é monótono.

Exemplo 6. (CRUZ NETO et al., 2002, Proposição 3.2)

Para todo $p \in M$ fixado, o campo vetorial gradiente $\text{grad } \frac{1}{2}d^2(p, q) = -\exp_p^{-1} q$ é fortemente monótono, onde $q \in M$.

Exemplo 7. (CRUZ NETO et al., 2002, Proposição 3.4)

(i) *A função $f : M \rightarrow \mathbb{R}$ é convexa se, e somente se, o campo vetorial $\text{grad } f$ é monótono.*

(ii) *A função $f : M \rightarrow \mathbb{R}$ é estritamente convexa se, e somente se, o campo vetorial $\text{grad } f$ é estritamente monótono.*

(iii) *A função $f : M \rightarrow \mathbb{R}$ é fortemente convexa se, e somente se, o campo vetorial $\text{grad } f$ é fortemente monótono.*

Observação 6. *Pelo Exemplo 6 o campo vetorial $\text{grad } \frac{1}{2}d^2(p, \cdot)$ é fortemente monótono.*

Então, pelo item (iii) do Exemplo 7 a função $\frac{1}{2}d^2(p, \cdot)$ é fortemente convexa.

Observação 7. *Os exemplos acima de campos vetoriais monótonos são subdiferenciais de funções convexas. Veremos no capítulo de experimentos numéricos exemplos de campos*

vetoriais monótonos que não são subdiferenciais de funções convexas, conforme (ROCKAFELLAR, 1966, Teorema 3), por que não são operadores ciclicamente monótonos; veja (PHELPS, 1997).

Exemplo 8. Suponha que $A \in \mathcal{X}(M)$ é monótono. Seja $z \in D(A)$ e defina o campo vetorial $B : M \rightarrow 2^{TM}$ por

$$B(x) := A(x) - \rho \exp_x^{-1} z, \quad \forall x \in D(A), \quad \rho > 0.$$

Então, B é fortemente monótono; veja (LI et al., 2009) e (LI et al., 2011). De fato, dados $x, y \in D(A)$, $u \in B(x)$ e $v \in B(y)$, temos que

$$u = \bar{u} - \rho \exp_x^{-1} z \quad e \quad v = \bar{v} - \rho \exp_y^{-1} z, \quad \bar{u} \in A(x), \bar{v} \in A(y).$$

Segue da monotonicidade de A que $\langle \bar{u}, \exp_x^{-1} y \rangle - \langle \bar{v}, -\exp_y^{-1} x \rangle \leq 0$, o que implica em

$$\langle u + \rho \exp_x^{-1} z, \exp_x^{-1} y \rangle - \langle v + \rho \exp_y^{-1} z, -\exp_y^{-1} x \rangle \leq 0.$$

Assim,

$$\langle u, \exp_x^{-1} y \rangle - \langle v, -\exp_y^{-1} x \rangle \leq -\rho[\langle \exp_x^{-1} z, \exp_x^{-1} y \rangle + \langle \exp_y^{-1} z, \exp_y^{-1} x \rangle]. \quad (1.7)$$

Agora, pelo triângulo geodésico $\Delta(z, x, y)$ com $\theta = \angle(\exp_x^{-1} z, \exp_x^{-1} y)$, então pelo Teorema 3, temos

$$d^2(z, x) + d^2(x, y) - 2\langle \exp_x^{-1} z, \exp_x^{-1} y \rangle \leq d^2(z, y). \quad (1.8)$$

Analogamente, considere o triângulo geodésico $\Delta(z, y, x)$ com $\theta = \angle(\exp_y^{-1} z, \exp_y^{-1} x)$, temos

$$d^2(z, y) + d^2(y, x) - 2\langle \exp_y^{-1} z, \exp_y^{-1} x \rangle \leq d^2(z, x). \quad (1.9)$$

Somando as Desigualdades (1.8) e (1.9), temos

$$2d^2(x, y) - 2[\langle \exp_x^{-1} z, \exp_x^{-1} y \rangle + \langle \exp_y^{-1} z, \exp_y^{-1} x \rangle] \leq 0,$$

isto é,

$$-[\langle \exp_x^{-1} z, \exp_x^{-1} y \rangle + \langle \exp_y^{-1} z, \exp_y^{-1} x \rangle] \leq -d^2(x, y). \quad (1.10)$$

Segue de (1.7) e (1.10) que

$$\langle u, \exp_x^{-1} y \rangle - \langle v, -\exp_y^{-1} x \rangle \leq -\rho d^2(x, y), \forall u \in B(x) \text{ e } \forall v \in B(y).$$

Portanto, B é fortemente monótono.

Exemplo 9. (BATISTA et al., 2016, Lema 3.1)

Sejam $A, B \in \mathcal{X}(M)$ monótonos maximais tais que $D(A) = D(B) = M$. Então, $A + B$ é um campo vetorial monótono maximal.

Pelo Exemplo 9 a soma de dois campos vetoriais monótonos é sempre monótono. No entanto, a diferença de dois campos vetoriais monótonos não é necessariamente monótono.

Exemplo 10. Sejam $A(X) = \text{grad } f_1(X)$ e $B(X) = \text{grad } f_2(X)$ os campos vetoriais gradientes das funções convexas $f_1(X) = a \ln(\det X)^2$ e $f_2(X) = b \ln(\det X)^4$, respectivamente, conforme Exemplo 4 nos itens (i) e (ii), respectivamente. Segue do Exemplo 7 que A e B são campos vetoriais monótonos. Por outro lado, para $B(X) - A(X) = \text{grad } f_3(X)$, onde $f_3(X) = f_2(X) - f_1(X)$ é uma função não convexa pelo Exemplo 4 item (iii), assim, pelo Exemplo 4, temos que $B - A$ não é monótono.

As noções de semicontinuidade superior e Kuratowski semicontinuidade superior, bem como os conceitos de localmente limitado foram estendidos para o cenário de variedades de Hadamard por (LI et al., 2009) na seguinte definição:

Definição 19. Sejam $A \in \mathcal{X}(M)$ e $\bar{x} \in D(A)$. Então, dizemos que A é

- (i) *semicontínua superior em \bar{x} se, para qualquer conjunto aberto V satisfazendo $A(\bar{x}) \subset V \subset T_{\bar{x}}M$, existe uma vizinhança aberta $U(\bar{x})$ de \bar{x} tal que $P_{\bar{x},x}A(x) \subset V$ para todo $x \in U(\bar{x})$;*
- (ii) *Kuratowski semicontínua superior em \bar{x} se, para qualquer sequência $\{x_k\} \subseteq D(A)$ e $\{u_k\} \subset TM$ com cada $u_k \in A(x_k)$, as relações $\lim_{k \rightarrow \infty} x_k = \bar{x}$ e $\lim_{k \rightarrow \infty} u_k = \bar{u}$ implicam $\bar{u} \in A(\bar{x})$.*
- (iii) *localmente limitado em \bar{x} se existe uma vizinhança aberta $U(\bar{x})$ de \bar{x} tal que o conjunto $\cup_{x \in U(\bar{x})} A(x)$ é limitado;*

(iv) *semicontínua superior (resp. Kuratowski semicontínuo superior, localmente limitado) em M se for semicontínua superior (resp. Kuratowski semicontínuo superior, localmente limitado) em cada ponto $\bar{x} \in D(A)$.*

Observação 8. *Como observado em (LI et al., 2009) semicontinuidade superior implica Kuratowski semicontínuo superior. Por outro lado, a recíproca também é verdadeira se A for localmente limitado em M .*

Vejamos em seguida um exemplo de campo vetorial monótono maximal e Kuratowski semicontínuo superior.

Exemplo 11. *(LI et al., 2009, Teorema 5.1)*

Seja $f : M \rightarrow \mathbb{R}$ uma função convexa semicontínua inferiormente em M . Então, ∂f é um campo vetorial monótono maximal e Kuratowski semicontínuo superior.

A seguinte proposição mostra que a maximalidade implica Kuratowski semicontínuo superior. Seja $\bar{x} \in D(A)$ e seja $T : T_{\bar{x}}M \rightarrow 2^{T_{\bar{x}}M}$ aplicação definida por

$$T(\mathbf{u}) = P_{\bar{x}, \exp_{\bar{x}}} A(\exp_{\bar{x}} \mathbf{u}), \quad \forall \mathbf{u} \in T_{\bar{x}}M. \quad (1.11)$$

Proposição 5. *Seja $A \in \mathcal{X}(M)$. Considere as seguintes afirmações.*

(i) *O campo vetorial A é monótono maximal.*

(ii) *Para cada $\bar{x} \in D(A)$, a aplicação $T : T_{\bar{x}}M \rightarrow 2^{T_{\bar{x}}M}$ definida em (1.11) é Kuratowski semicontínuo superior em $T_{\bar{x}}M$.*

(iii) *O campo vetorial A é Kuratowski semicontínuo superior em M .*

Então, (i) \Rightarrow (ii) \Rightarrow (iii).

Demonstração. Veja (LI et al., 2009, Proposição 3.5). □

Em seguida, apresentamos alguns resultados de campos vetoriais monótonos maximais que serão usados posteriormente na prova de nossos resultados.

Proposição 6. *Suponha que $A \in \mathcal{X}(M)$ é monótono maximal e $D(A) = M$. Então, A é localmente limitado em M .*

Demonstração. Veja (LI et al., 2009, Lema 3.6). □

Exemplo 12. *Seja $f : M \rightarrow \mathbb{R}$ uma função convexa semicontínua inferiormente em M . Segue do Exemplo 11 e Proposição 6 que ∂f é localmente limitado.*

Observação 9. *Suponha que $A \in \mathcal{X}(M)$ seja monótono maximal e $D(A) = M$. Se $\{x^k\}$ é limitada e $v^k \in A(x^k)$ para todo $k \in \mathbb{N}$, então $\{v^k\}$ é limitada. De fato, se $A = \partial f$ onde f é como no Exemplo 12, a afirmação segue pela Proposição 2. Agora se A é um campo vetorial monótono maximal qualquer, segue da limitação de $\{x^k\}$ que a sequência $\{d(x^n, x^k)\}$ é limitada para todo $n, k \in \mathbb{N}$. Assim existe $C > 0$ tal que $d(x^0, x^k) \leq C$, para todo $k \in \mathbb{N}$. Defina a vizinhança aberta de x^0 dada por $V = \{x \in M; d(x^0, x) < C + 1\}$. Segue da Proposição 6 e Definição 19 (iii) que $\cup_{x \in V} A(x)$ é limitado. Como $x^k \in V$ e $v^k \in A(x^k)$, concluímos que $\{v^k\}$ é limitada e a afirmação está provada. Esta afirmação pode ser encontrada também em (BATISTA et al., 2016, Proposição 3.6).*

Dizemos que $x \in D(A)$ é uma singularidade de A se $0 \in A(x)$. O conjunto de todas as singularidades de A é denotado por $A^{-1}(0)$, isto é,

$$A^{-1}(0) := \{x \in D(A); 0 \in A(x)\}.$$

A seguinte proposição sobre a unicidade de singularidade é uma consequência direta da monotonicidade estrita.

Proposição 7. *Seja $A \in \mathcal{X}(M)$ estritamente monótono. Então, A tem no máximo uma singularidade.*

Proposição 8. *Seja $A \in \mathcal{X}(M)$ um campo vetorial maximal, fortemente monótono com o domínio $D(A) = M$. Então, existe uma única singularidade de A .*

Demonstração. Veja (LI et al., 2009, Teorema 4.3). □

O conceito de resolvente de um campo vetorial $A \in \mathcal{X}(M)$ foi introduzido e estudado para o cenário de variedades de Hadamard em (LI et al., 2011) como veremos na seguinte definição e proposições posteriores.

Definição 20. *Dados $\lambda > 0$ e $A \in \mathcal{X}(M)$, o resolvente de A de ordem λ é a aplicação de valores múltiplos $J_\lambda^A : M \rightarrow 2^M$ definida por*

$$J_\lambda^A(x) := \{z \in M; x \in \exp_z \lambda A(z)\}, \quad \forall x \in M.$$

Observação 10. Como mencionado em (LI et al., 2011, Observação 3) pela definição de resolvente de um campo vetorial, o domínio do resolvente J_λ^A é a imagem do campo vetorial definido por $x \mapsto \exp_x \lambda A(x)$. Vamos denotar essa imagem por $\mathcal{R}(\exp_{(\cdot)} \lambda A(\cdot))$. Então temos que

$$D(J_\lambda^A) = \mathcal{R}(\exp_{(\cdot)} \lambda A(\cdot)).$$

Apresentaremos agora a definição de aplicações firmemente não expansivas em variedades de Hadamard, introduzida nesse cenário em (LI et al., 2011).

Definição 21. Seja $C \subset M$ um conjunto convexo e fechado. Uma aplicação $T : C \rightarrow M$ é firmemente não expansiva se a aplicação $\phi : [0, 1] \rightarrow \mathbb{R}_+$ definida por

$$\phi(t) := d(\exp_x t \exp_x^{-1} Tx, \exp_y t \exp_y^{-1} Ty), \quad \forall t \in [0, 1]$$

é não crescente para quaisquer $x, y \in C$.

Em seguida, apresentamos um resultado importante em aplicações firmemente não expansivas.

Proposição 9. Seja $T : C \subseteq M \rightarrow M$, onde C é um conjunto convexo e fechado. As seguintes afirmações são equivalentes.

- (i) T é firmemente não expansiva;
- (ii) Para quaisquer $x, y \in C$ e $t \in [0, 1]$, tem-se

$$d(T(x), T(y)) \leq d(\exp_x t \exp_x^{-1} Tx, \exp_y t \exp_y^{-1} Ty).$$

Demonstração. Veja (LI et al., 2011, Proposição 5). □

O seguinte resultado estabelece uma relação da propriedade de não expansividade do resolvente com a monotonicidade do campo vetorial correspondente.

Proposição 10. Seja $A \in \mathcal{X}(M)$. As seguintes afirmações são válidas para qualquer $\lambda > 0$.

- (i) O campo vetorial A é monótono se, e somente se, J_λ^A é de valor único e firmemente não expansivo.

(ii) Se $D(A) = M$, o campo vetorial A é monótono maximal se, e somente se, J_λ^A é de valor único, firmemente não expansivo e o domínio $D(J_\lambda^A) = M$.

Demonstração. Veja (LI et al., 2011, Teorema 4). □

Capítulo 2

Método do ponto proximal

Os resultados deste capítulo encontram-se em (ANDRADE et al., 2021).

Considere uma classe especial de problemas de minimização não convexos, conhecidos como problemas de minimização DC:

$$\min_{x \in \mathbb{R}^n} f(x) = \min_{x \in \mathbb{R}^n} g(x) - h(x), \quad (2.1)$$

onde $g, h : \mathbb{R}^n \rightarrow \mathbb{R}$ são funções convexas. A função f é chamada uma função DC (i.e. diferença de duas funções convexas) e é limitada inferiormente.

O estudo de algoritmos para resolver problemas de minimização DC tem seu primeiro e mais popular algoritmo, denominado DCA (DC Algoritmos), proposto e apresentado em (PHAM; SOUAD, 1986) no ano de 1986. Desde este algoritmo, várias variantes de DCA surgiram e várias questões teóricas e práticas foram descobertas ao longo dos anos, resultando em uma vasta literatura sobre o assunto; veja (THI; PHAM, 2018) para uma perspectiva histórica.

No ano de 2003, (SUN et al., 2003) propuseram um algoritmo de ponto proximal para o Problema (2.1) que usam as propriedades das duas funções convexas separadamente. Em 2006, esse trabalho foi complementado com os resultados de (MOUDAFI; MAINGÉ, 2006).

Em (SOUZA; OLIVEIRA, 2015) o algoritmo proposto por (SUN et al., 2003) é estendido à estrutura das variedades Riemannianas, onde é proposto um algoritmo para encontrar pontos críticos da função objetivo f , ou seja, resolver o seguinte problema

$$\text{encontrar } x^* \in M \text{ tal que } \partial g(x^*) \cap \partial h(x^*) \neq \emptyset, \quad (2.2)$$

onde M é uma variedade de Hadamard, ∂g e ∂h são campos vetoriais de valores múltiplos monótonos maximais. O algoritmo é o seguinte:

Algoritmo DCPPA.

Passo 1: Dado um ponto inicial $x^0 \in M$ e uma sequência de números positivos $\{\mu_k\}$ tal que $0 < a \leq \mu_k \leq b$.

Passo 2: Dado $x^k \in M$, calcule

$$w^k \in \partial h(x^k) \text{ e defina } y^k := \exp_{x^k}(\mu_k w^k).$$

Passo 3: Calcule

$$x^{k+1} := \arg \min_{x \in M} \{g(x) + \frac{1}{2\mu_k} d^2(x, y^k)\}.$$

Se $x^{k+1} = x^k$, pare. Caso contrário, defina $k := k + 1$ e retorne ao Passo 2.

Neste trabalho, estamos interessados em resolver o seguinte problema:

$$\text{encontrar } x^* \in M \text{ tal que } A(x^*) \cap B(x^*) \neq \emptyset \tag{2.3}$$

que é uma generalização do Problema (2.2) a campos vetoriais de valores múltiplos monótonos maximais quaisquer A e B no contexto de variedades de Hadamard. Mais precisamente, apresentamos aqui métodos que generalizam o DCPPA proposto por (SOUZA; OLIVEIRA, 2015). No cenário linear, (MOUDAFI, 2015) propôs um algoritmo baseado no operador resolvente e na aproximação de Yosida para resolver (2.3). Vale ressaltar que no cenário linear nossos métodos são diferentes dos métodos propostos em (MOUDAFI, 2008, 2015).

2.1 Algoritmo MPP

Ao longo deste texto, assumimos que M é uma variedade de Hadamard de dimensão finita e $A, B \in \mathcal{X}(M)$ são campos vetoriais monótonos maximais com $D(A) = D(B) = M$.

Definição 22. *Um ponto $x^* \in M$ satisfazendo (2.3), diz-se um ponto crítico do campo vetorial possivelmente não monótono $A - B$. O conjunto dos pontos críticos de $A - B$ é definido por*

$$S = \{x^* \in M; A(x^*) \cap B(x^*) \neq \emptyset\}.$$

Observação 11. *No caso particular, onde $A = \partial g$ e $B = \partial h$ são subdiferenciais das funções convexas g e h respectivamente, S é o conjunto solução do Problema (2.2), isto é, o conjunto S coincide com o conjunto dos pontos críticos da função DC f . Veja (JOKI et al., 2018) para uma discussão muito interessante sobre a criticidade das funções DC.*

Neste trabalho, supomos que $S \neq \emptyset$. Note que uma condição necessária e suficiente para que um ponto x seja um ponto crítico de $A - B$, ou seja, solução de (2.3) é que

$$\frac{1}{\mu} \exp_x^{-1} \mathbf{y} \in A(x),$$

onde $\mathbf{y} = \exp_x(\mu w)$, para qualquer $w \in B(x)$ e $\mu > 0$ um número real. Baseado na afirmação acima vamos definir nosso método do ponto proximal (MPP) para resolver o Problema (2.3).

Método do ponto proximal - MPP.

Passo 1: Dado um ponto inicial $x^0 \in M$, considere $\mathbf{a}, \mathbf{b} > 0$. Seja $\{\mu_k\}$ uma sequência auxiliar de números positivos tal que

$$\mathbf{a} \leq \mu_k \leq \mathbf{b}. \tag{2.4}$$

Passo 2: Dado $x^k \in M$, escolha

$$w^k \in B(x^k) \text{ e defina } \mathbf{y}^k := \exp_{x^k}(\mu_k w^k). \tag{2.5}$$

Passo 3: Calcular $x^{k+1} \in M$ tal que

$$0 \in A(x^{k+1}) - \frac{1}{\mu_k} \exp_{x^{k+1}}^{-1} \mathbf{y}^k. \tag{2.6}$$

Se $x^{k+1} = x^k$, pare. Caso contrário, defina $k := k + 1$ e retorne ao Passo 2.

Observação 12. *Note que se $B(x) = 0$, então por (2.5) teremos $\mathbf{y}^k = x^k$ tornando (2.6) em*

$$0 \in A(x^{k+1}) - \frac{1}{\mu_k} \exp_{x^{k+1}}^{-1} x^k,$$

recuperando assim o algoritmo proposto por (LI et al., 2009) para o problema de encontrar singularidades de um campo vetorial de valores múltiplos monótono maximal em variedades de Hadamard, que generaliza o método do ponto proximal para resolver problemas de

minimização convexa em variedades de Hadamard considerado por (FERREIRA; OLIVEIRA, 2002). Se $A = \partial g$ e $B = \partial h$, onde as aplicações ∂g e ∂h são subdiferenciais das funções convexas g e h , respectivamente, então usando a Proposição 4, o MPP se torna o algoritmo

$$w^k \in \partial h(x^k), \quad y^k := \exp_{x^k}(\mu_k w^k).$$

e

$$0 \in \partial g(x^{k+1}) - \frac{1}{\mu_k} \exp_{x^{k+1}}^{-1} y^k \Rightarrow x^{k+1} = \arg \min_{x \in M} \{g(x) + \frac{1}{2\mu_k} d^2(x, y^k)\}$$

proposto por (SOUZA; OLIVEIRA, 2015) para resolver o problema de otimização DC. Portanto, o MPP generaliza alguns métodos existentes em variedades de Hadamard para um contexto mais geral. Além disso, alguns resultados apresentados neste trabalho são novos inclusive no cenário euclidiano.

Em seguida, estabelecemos a boa definição das sequências $\{x^k\}$ e $\{y^k\}$, bem como uma regra prática de parada para o MPP.

Proposição 11. *As seguintes afirmações são válidas:*

- (i) A sequência $\{x^k\}$ e $\{y^k\}$ são bem definidas;
- (ii) Se $x^{k+1} = x^k$, então MPP para em uma solução de (2.3).

Demonstração. A boa definição das sequências $\{y^k\}$ and $\{x^k\}$ segue diretamente do fato que M é uma variedade de Hadamard e, para cada $k \in \mathbb{N}$ o campo vetorial

$$x \mapsto A(x) - \frac{1}{\mu_k} \exp_x^{-1} y^k$$

é fortemente monótono pelo Exemplo 8 e monótono maximal pelos Exemplos 9 e 11 e, portanto, tem uma única singularidade pela Proposição 8.

Para provar a segunda afirmação, observe que a partir de (2.5) e (2.6), temos que

$$\frac{1}{\mu_k} \exp_{x^k}^{-1} y^k \in B(x^k) \quad \text{e} \quad \frac{1}{\mu_k} \exp_{x^{k+1}}^{-1} y^k \in A(x^{k+1}).$$

Assim, se $x^{k+1} = x^k$, obtemos

$$\frac{1}{\mu_k} \exp_{x^k}^{-1} y^k \in B(x^k) \quad \text{e} \quad \frac{1}{\mu_k} \exp_{x^k}^{-1} y^k \in A(x^k).$$

Portanto, $\frac{1}{\mu_k} \exp_{x^k}^{-1} \mathbf{y}^k \in A(x^k) \cap B(x^k)$, isto é, $A(x^k) \cap B(x^k) \neq \emptyset$ e conseqüentemente x^k é uma solução de (2.3). \square

2.2 Análise de convergência

O objetivo desta seção é estudar as propriedades de convergência do MPP. A partir de agora, consideramos $\{x^k\}$ a sequência gerada pelo MPP e, tendo em vista a Proposição 11, assumimos que $x^{k+1} \neq x^k$ para todos os $k \in \mathbb{N}$. Agora vamos estabelecer suas propriedades de convergência. Começando com a seguinte proposição que desempenha um papel importante na análise de convergência do MPP.

Proposição 12. *Suponha que A ou B é ρ -fortemente monótono. Então,*

$$\left(\rho + \frac{1}{b}\right) d(x^{k+1}, x^k) \leq \|P_{x^{k+1}, x^k} v^k - u^{k+1}\|, \quad (2.7)$$

para qualquer $v^k \in A(x^k)$ e $u^{k+1} \in B(x^{k+1})$.

Demonstração. Suponhamos que A é monótono e B é ρ -fortemente monótono. O outro caso é análogo. Por (2.5) e (2.6), temos

$$\frac{1}{\mu_k} \exp_{x^k}^{-1} \mathbf{y}^k \in B(x^k) \quad \text{e} \quad \frac{1}{\mu_k} \exp_{x^{k+1}}^{-1} \mathbf{y}^k \in A(x^{k+1}).$$

Dado $x \in M$, sendo B ρ -fortemente monótono, obtemos

$$\left\langle \frac{1}{\mu_k} \exp_{x^k}^{-1} \mathbf{y}^k, \exp_{x^k}^{-1} x \right\rangle - \langle u, -\exp_x^{-1} x^k \rangle \leq -\rho d^2(x, x^k), \quad \forall u \in B(x). \quad (2.8)$$

Segue da monotonicidade de A que

$$\left\langle \frac{1}{\mu_k} \exp_{x^{k+1}}^{-1} \mathbf{y}^k, \exp_{x^{k+1}}^{-1} x \right\rangle \leq \langle v, -\exp_x^{-1} x^{k+1} \rangle, \quad \forall v \in A(x). \quad (2.9)$$

Tomando $x = x^{k+1}$ em (2.8), $x = x^k$ em (2.9) e somando essas desigualdades, temos que

$$\begin{aligned} & \frac{1}{\mu_k} [\langle \exp_{x^k}^{-1} \mathbf{y}^k, \exp_{x^k}^{-1} x^{k+1} \rangle + \langle \exp_{x^{k+1}}^{-1} \mathbf{y}^k, \exp_{x^{k+1}}^{-1} x^k \rangle] \leq \\ & \langle u^{k+1}, -\exp_{x^{k+1}}^{-1} x^k \rangle + \langle v^k, -\exp_{x^k}^{-1} x^{k+1} \rangle - \rho d^2(x^{k+1}, x^k), \end{aligned} \quad (2.10)$$

para qualquer $\mathbf{v}^k \in A(\mathbf{x}^k)$ e $\mathbf{u}^{k+1} \in B(\mathbf{x}^{k+1})$. Seja $\Delta(\mathbf{y}^k, \mathbf{x}^k, \mathbf{x}^{k+1})$ o triângulo geodésico com $\theta = \angle(\exp_{\mathbf{x}^k}^{-1} \mathbf{y}^k, \exp_{\mathbf{x}^k}^{-1} \mathbf{x}^{k+1})$. Então, pelo Teorema 3, temos que

$$\mathbf{d}^2(\mathbf{y}^k, \mathbf{x}^k) + \mathbf{d}^2(\mathbf{x}^k, \mathbf{x}^{k+1}) - 2\langle \exp_{\mathbf{x}^k}^{-1} \mathbf{y}^k, \exp_{\mathbf{x}^k}^{-1} \mathbf{x}^{k+1} \rangle \leq \mathbf{d}^2(\mathbf{y}^k, \mathbf{x}^{k+1}). \quad (2.11)$$

De forma análoga, considerando o triângulo geodésico $\Delta(\mathbf{y}^k, \mathbf{x}^{k+1}, \mathbf{x}^k)$ com

$$\theta = \angle(\exp_{\mathbf{x}^{k+1}}^{-1} \mathbf{y}^k, \exp_{\mathbf{x}^{k+1}}^{-1} \mathbf{x}^k),$$

temos

$$\mathbf{d}^2(\mathbf{y}^k, \mathbf{x}^{k+1}) + \mathbf{d}^2(\mathbf{x}^{k+1}, \mathbf{x}^k) - 2\langle \exp_{\mathbf{x}^{k+1}}^{-1} \mathbf{y}^k, \exp_{\mathbf{x}^{k+1}}^{-1} \mathbf{x}^k \rangle \leq \mathbf{d}^2(\mathbf{y}^k, \mathbf{x}^k). \quad (2.12)$$

Adicionando (2.11) e (2.12), obtemos

$$\mathbf{d}^2(\mathbf{x}^{k+1}, \mathbf{x}^k) \leq \langle \exp_{\mathbf{x}^k}^{-1} \mathbf{y}^k, \exp_{\mathbf{x}^k}^{-1} \mathbf{x}^{k+1} \rangle + \langle \exp_{\mathbf{x}^{k+1}}^{-1} \mathbf{y}^k, \exp_{\mathbf{x}^{k+1}}^{-1} \mathbf{x}^k \rangle. \quad (2.13)$$

Sendo $\mu_k > 0$, para todo $k \in \mathbb{N}$, e usando (2.13) in (2.10), temos

$$\frac{1}{\mu_k} \mathbf{d}^2(\mathbf{x}^{k+1}, \mathbf{x}^k) \leq \langle \mathbf{u}^{k+1}, -\exp_{\mathbf{x}^{k+1}}^{-1} \mathbf{x}^k \rangle + \langle \mathbf{v}^k, -\exp_{\mathbf{x}^k}^{-1} \mathbf{x}^{k+1} \rangle - \rho \mathbf{d}^2(\mathbf{x}^{k+1}, \mathbf{x}^k). \quad (2.14)$$

Assim, usando o fato que $0 < \mathbf{a} \leq \mu_k \leq \mathbf{b}$, obtemos

$$\left(\rho + \frac{1}{\mathbf{b}} \right) \mathbf{d}^2(\mathbf{x}^{k+1}, \mathbf{x}^k) \leq \langle \mathbf{u}^{k+1}, -\exp_{\mathbf{x}^{k+1}}^{-1} \mathbf{x}^k \rangle + \langle \mathbf{v}^k, -\exp_{\mathbf{x}^k}^{-1} \mathbf{x}^{k+1} \rangle, \quad (2.15)$$

para todo $\mathbf{v}^k \in A(\mathbf{x}^k)$ e $\mathbf{u}^{k+1} \in B(\mathbf{x}^{k+1})$. Por outro lado, usando as propriedades de transporte paralelo, temos que $\exp_{\mathbf{x}^{k+1}}^{-1} \mathbf{x}^k = -\mathbf{P}_{\mathbf{x}^{k+1}, \mathbf{x}^k} \exp_{\mathbf{x}^k}^{-1} \mathbf{x}^{k+1}$ (veja (LI et al., 2009)), e consequentemente

$$\langle \mathbf{u}^{k+1}, -\exp_{\mathbf{x}^{k+1}}^{-1} \mathbf{x}^k \rangle + \langle \mathbf{v}^k, -\exp_{\mathbf{x}^k}^{-1} \mathbf{x}^{k+1} \rangle = \langle \mathbf{P}_{\mathbf{x}^{k+1}, \mathbf{x}^k} \mathbf{v}^k - \mathbf{u}^{k+1}, \exp_{\mathbf{x}^{k+1}}^{-1} \mathbf{x}^k \rangle.$$

Assim, usando este fato em (2.15), temos

$$\left(\rho + \frac{1}{\mathbf{b}} \right) \mathbf{d}^2(\mathbf{x}^{k+1}, \mathbf{x}^k) \leq \langle \mathbf{P}_{\mathbf{x}^{k+1}, \mathbf{x}^k} \mathbf{v}^k - \mathbf{u}^{k+1}, \exp_{\mathbf{x}^{k+1}}^{-1} \mathbf{x}^k \rangle, \quad (2.16)$$

para todo $\mathbf{v}^k \in A(\mathbf{x}^k)$ e $\mathbf{u}^{k+1} \in B(\mathbf{x}^{k+1})$. Aplicando a desigualdade de Cauchy-Schwarz em (2.16), obtemos

$$\left(\rho + \frac{1}{b}\right) d^2(\mathbf{x}^{k+1}, \mathbf{x}^k) \leq \|P_{\mathbf{x}^{k+1}, \mathbf{x}^k} \mathbf{v}^k - \mathbf{u}^{k+1}\| \|\exp_{\mathbf{x}^{k+1}}^{-1} \mathbf{x}^k\|,$$

para todo $\mathbf{v}^k \in A(\mathbf{x}^k)$ e $\mathbf{u}^{k+1} \in B(\mathbf{x}^{k+1})$. Portanto, usando agora na desigualdade acima o fato que, em variedades de Hadamard, $\|\exp_{\mathbf{x}^{k+1}}^{-1} \mathbf{x}^k\| = d(\mathbf{x}^{k+1}, \mathbf{x}^k)$, segue o resultado desejado. \square

Observação 13. *Vale ressaltar que supor a monotonicidade forte de A ou B em vez de apenas monotonicidade não é uma suposição restritiva. De fato, se a monotonicidade forte não vale para A e B , podemos obter outra decomposição que satisfaça esta condição da seguinte forma:*

$$T(\mathbf{x}) = A(\mathbf{x}) - B(\mathbf{x}) = [A(\mathbf{x}) - \rho \exp_{\mathbf{x}}^{-1} \mathbf{y}] - [B(\mathbf{x}) - \rho \exp_{\mathbf{x}}^{-1} \mathbf{y}],$$

para $\rho > 0$ arbitrário e $\mathbf{y} \in M$ fixado. Sabemos que $\tilde{A}(\mathbf{x}) = A(\mathbf{x}) - \rho \exp_{\mathbf{x}}^{-1} \mathbf{y}$ e $\tilde{B}(\mathbf{x}) = B(\mathbf{x}) - \rho \exp_{\mathbf{x}}^{-1} \mathbf{y}$ são ρ -fortemente monótonos; veja Exemplo 8.

A seguir mostramos que a distância Riemanniana entre duas iteradas consecutivas do MPP tende a zero quando k tende a ∞ , que é uma propriedade clássica de métodos de ponto proximal.

Proposição 13. *Suponha que $\{\mathbf{x}^k\}$ é limitada, então existem constantes $L, M \geq 0$ tais que $\|P_{\mathbf{x}^{k+1}, \mathbf{x}^k} \mathbf{v}^k - \mathbf{u}^{k+1}\| \leq L$ e $\limsup_{k \rightarrow \infty} d(\mathbf{x}^{k+1}, \mathbf{x}^k) = M$, para quaisquer $\mathbf{v}^k \in A(\mathbf{x}^k)$, $\mathbf{u}^{k+1} \in B(\mathbf{x}^{k+1})$ para todo k . Além disso, se A ou B é ρ -fortemente monótono e*

$$\left(\rho + \frac{1}{b}\right) M > L, \tag{2.17}$$

então $\lim_{k \rightarrow \infty} d(\mathbf{x}^{k+1}, \mathbf{x}^k) = 0$.

Demonstração. Como $\{\mathbf{x}^k\}$ é limitado, segue da maximalidade de A e B (veja Proposição 6) que existem constantes $K_1, K_2 > 0$ tais que $\|\mathbf{v}^k\| < K_1$ e $\|\mathbf{u}^k\| < K_2$ para todo k . Assim, pela desigualdade triangular, temos

$$\|P_{\mathbf{x}^{k+1}, \mathbf{x}^k} \mathbf{v}^k - \mathbf{u}^{k+1}\| \leq \|P_{\mathbf{x}^{k+1}, \mathbf{x}^k} \mathbf{v}^k\| + \|\mathbf{u}^{k+1}\|.$$

Usando o fato de que a aplicação transporte paralelo é uma isometria, temos

$$\|P_{x^{k+1}, x^k} v^k - u^{k+1}\| < L := K_1 + K_2, \quad (2.18)$$

para todo k , e conseqüentemente, a desigualdade está provada. Agora, como $\{x^k\}$ é limitada, segue que a sequência $\{d(x^{k+1}, x^k)\}$ é limitada. Denote por $M = \limsup_{k \rightarrow \infty} d(x^{k+1}, x^k)$. Assuma por contradição que $M > 0$ e, suponha que A ou B é ρ -fortemente monótono com $\rho > 0$ satisfazendo (2.17). Assim, combinando novamente (2.7) e (2.18), temos

$$\left(\rho + \frac{1}{b}\right) d(x^{k+1}, x^k) \leq \|P_{x^{k+1}, x^k} v^k - u^{k+1}\| \leq L.$$

Tomando o $\limsup_{k \rightarrow \infty}$ na desigualdade acima, isso implica que $(\rho + \frac{1}{b}) M \leq L$ que é uma contradição. Portanto, a afirmação está provada. \square

Observação 14. *Se $A(\cdot) = \partial g(\cdot)$ e $B(\cdot) = \partial h(\cdot)$ com g e h funções convexas e, conseqüentemente A e B são campos vetoriais monótonos maximais. Sendo $f(x) = g(x) - h(x)$, então (2.17) é substituída pela limitação inferior de f . No entanto, a prova ainda segue de uma desigualdade como (2.7), mais precisamente, segue da seguinte desigualdade $\frac{1}{b} d^2(x^{k+1}, x^k) \leq f(x^0) - f^*$, onde $f^* = \inf_{x \in M} f(x)$; veja (SOUZA; OLIVEIRA, 2015, Proposição 5). Lidando com campos vetoriais monótonos maximais gerais, precisamos de uma suposição mais geral. Porém, vale a pena mencionar que a desigualdade (2.17) não é restritiva devido ao fato de que o parâmetro ρ pode ser tomado grande o suficiente para que (2.17) seja válida; veja Observação 13.*

2.2.1 Análise de convergência parcial

Ao lidar com funções DC, sabe-se que apenas resultados de convergência parcial são obtidos, no sentido de que os pontos de acumulação da sequência (se houver) são pontos críticos da função objetivo. Isso foi comprovado por vários métodos, por exemplo, DCA ver (PHAM; SOUAD, 1986; TAO; AN, 1997), DCA reforçado (ver (ARTACHO et al., 2018; ARTACHO; VUONG, 2020)), subgradiente escalonado (ver (FERREIRA et al., 2021)), método do ponto proximal (consulte (ALMEIDA et al., 2020; CRUZ NETO et al., 2020; MOUDAFI; MAINGÉ, 2006; SOUZA; OLIVEIRA, 2015; SOUZA et al., 2016; SUN et al., 2003)), método do feixe proximal (consulte (OLIVEIRA, 2019)), feixe duplo (con-

sulte (JOKI et al., 2018)), codiferencial (consulte (BAGIROV; UGON, 2011)) e método inercial (ver (OLIVEIRA; TCHEOU, 2019)). Para obter resultados de convergência total, ou seja, convergência de toda a sequência, alguns trabalhos têm considerado a propriedade Kurdyka-Lojasiewicz da função objetivo; ver por exemplo (ARTACHO; VUONG, 2020; CRUZ NETO et al., 2020; FERREIRA et al., 2021; THI et al., 2018).

Nesta subseção apresentamos resultados de convergência parcial para o MPP. Na subseção seguinte, apresentamos resultados de convergência total para o MPP para o caso mais geral da diferença de campos vetoriais monótonos maximais em variedades de Hadamard que estende, por exemplo, alguns resultados em (SOUZA; OLIVEIRA, 2015). Introduzimos uma suposição sobre o parâmetro proximal $\{\mu_k\}$ e uma continuidade de Lipschitz de B para obter a convergência total da sequência para um ponto crítico de $A - B$.

Teorema 6. *Suponha que $\{x^k\}$ é limitada e A ou B é ρ -fortemente monótono. Então, todo ponto de acumulação de $\{x^k\}$ é um ponto crítico de $A - B$.*

Demonstração. Considere x , um ponto de acumulação de $\{x^k\}$, e seja $\{x^{k_j}\}$ subsequência de $\{x^k\}$ convergindo para x . Sendo $\{\mu_k\}$ e $\{y^k\}$ limitadas, podemos tomar sem perda de generalidade subsequências $\{\mu_{k_j}\}$ e $\{y^{k_j}\}$ convergindo respectivamente para μ e y . Da Proposição 13, temos que $x^{k_j+1} \rightarrow x$. Segue por (2.5) que

$$\frac{1}{\mu_{k_j}} \exp_{x^{k_j}}^{-1} y^{k_j} \in B(x^{k_j}).$$

Pela monotonicidade de B , temos

$$\left\langle \frac{1}{\mu_{k_j}} \exp_{x^{k_j}}^{-1} y^{k_j}, \exp_{x^{k_j}}^{-1} z \right\rangle \leq \langle u, -\exp_z^{-1} x^{k_j} \rangle, \quad \forall u \in B(z), z \in M.$$

Fazendo $j \rightarrow +\infty$ na última desigualdade e usando a Proposição 1, obtemos

$$\left\langle \frac{1}{\mu} \exp_x^{-1} y, \exp_x^{-1} z \right\rangle \leq \langle u, -\exp_z^{-1} x \rangle, \quad \forall u \in B(z), z \in M,$$

e conseqüentemente, pela maximalidade de B , temos

$$\frac{1}{\mu} \exp_x^{-1} y \in B(x). \tag{2.19}$$

Agora, de (2.6), temos

$$\frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} \in \mathbf{A}(x^{k_j+1}).$$

De modo análogo, pela monotonicidade de \mathbf{A} , segue que

$$\left\langle \frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j}, \exp_{x^{k_j+1}}^{-1} \mathbf{z} \right\rangle \leq \langle \mathbf{u}, -\exp_z^{-1} x^{k_j+1} \rangle, \quad \forall \mathbf{u} \in \mathbf{A}(z), \forall z \in M,$$

e fazendo $j \rightarrow +\infty$, temos pela maximalidade de \mathbf{A} que

$$\frac{1}{\mu} \exp_x^{-1} \mathbf{y} \in \mathbf{A}(x). \quad (2.20)$$

Portanto, de (2.19) e (2.20), temos que $\frac{1}{\mu} \exp_x^{-1} \mathbf{y} \in \mathbf{A}(x) \cap \mathbf{B}(x)$ o que significa que (2.3) vale, ou seja, x é um ponto crítico de $\mathbf{A} - \mathbf{B}$. \square

Observação 15. *Vale ressaltar que a suposição de ρ -forte monotonicidade de \mathbf{A} ou \mathbf{B} foi usada na prova do Teorema 6 apenas para garantir que se $x^{k_j} \rightarrow x$ com $j \rightarrow \infty$, então $x^{k_j+1} \rightarrow x$. Em outras palavras, a hipótese acima foi assumido para aplicar Proposição 13, ou seja, para obter que $\lim_{k \rightarrow \infty} d(x^{k+1}, x^k) = 0$. Portanto, podemos substituir a ρ -forte monotonicidade de \mathbf{A} ou \mathbf{B} pela suposição de que $\lim_{k \rightarrow \infty} d(x^{k+1}, x^k) = 0$. Por exemplo, no caso monótono considerado em (LI et al., 2009) e no contexto de funções DC (possivelmente não convexas) estudadas em (SOUZA; OLIVEIRA, 2015) a condição $\lim_{k \rightarrow \infty} d(x^{k+1}, x^k) = 0$ é válida. Além disso, este tipo de condição é assumida, por exemplo, em Leustean et al. (LEUȘTEAN et al., 2018) que estudam o método do ponto proximal em espaços $CAT(0)$ (incluindo variedades de Hadamard); veja Condição (C2) e Lema 3.3.*

2.2.2 Condições suficientes para limitação

Como mencionado anteriormente, alguns trabalhos vêm provando a convergência total da sequência para diversos métodos aplicados a funções DC supondo a propriedade Kurdyka-Lojasiewicz da função objetivo; veja (ARTACHO; VUONG, 2020; CRUZ NETO et al., 2020; FERREIRA et al., 2021; THI et al., 2018). No entanto, mesmo neste caso, supõe-se que a sequência gerada pelo método seja limitada. O objetivo desta subseção é introduzir condições suficientes para garantir a limitação da sequência gerada pelo nosso método. Vale ressaltar que este resultado é novo mesmo para problemas DC no cenário

Euclidiano. Para isso, consideramos as seguintes suposições no campo vetorial \mathbf{B} e nos parâmetros proximal μ_k . A partir de agora, vamos considerar o MPP substituindo a condição (2.4) pela condição mais fraca:

$$0 < \mu_k \leq \mathbf{b}. \quad (2.21)$$

(A1) Dados $\mathbf{x}, \mathbf{y} \in \mathbf{M}$, existe uma constante $L > 0$ tal que $\|\mathbf{u} - \mathbf{P}_{\mathbf{x}, \mathbf{y}} \mathbf{v}\| \leq Ld(\mathbf{x}, \mathbf{y})$, para quaisquer $\mathbf{u} \in \mathbf{B}(\mathbf{x})$ e $\mathbf{v} \in \mathbf{B}(\mathbf{y})$.

(A2) A sequência de números positivos $\{\mu_k\}$ satisfaz $\sum_{k=0}^{\infty} \mu_k < \infty$.

Observação 16. *Vale a pena mencionar que a suposição (A1) é uma extensão natural da continuidade Lipschitz do gradiente da função DC para a diferença de campos vetoriais monótonos maximais. Na literatura de algoritmo para funções DC alguns trabalhos têm considerado o caso não suave onde $f(\mathbf{x}) = g(\mathbf{x}) - h(\mathbf{x})$ com g, h funções convexas, possivelmente não suaves e h uma função $C^{1,1}$, ou seja, h é diferenciável e seu gradiente é Lipschitz contínuo; veja por exemplo (ARTACHO et al., 2018; ARTACHO; VUONG, 2020; CRUZ NETO et al., 2020; THI et al., 2018; SOUZA et al., 2016). No contexto de campos vetoriais, a condição (A1) implica que o campo vetorial \mathbf{B} é ponto-ponto. Por outro lado, a suposição (A2) não é restritiva, pois o parâmetro μ_k é escolhido livremente satisfazendo $0 < \mu_k \leq \mathbf{b}$, para todo $k \in \mathbb{N}$; veja por exemplo (FERREIRA; OLIVEIRA, 2002; MOUDAFI, 2015)*

Por uma questão de completude, apresentamos a seguinte ferramenta técnica que será usada posteriormente.

Lema 1. *Sejam $\{\alpha_k\}$, $\{\beta_k\}$ e $\{\gamma_k\}$ três sequências de números não negativos satisfazendo $\alpha_{k+1} \leq (1 + \beta_k)\alpha_k + \gamma_k$. Se $\sum_{k=0}^{\infty} \beta_k < +\infty$ e $\sum_{k=0}^{\infty} \gamma_k < +\infty$, então $\{\alpha_k\}$ é convergente.*

Demonstração. Veja (POLYAK, 1987, Lema 2 página 44). □

Sob as afirmações (A1) e (A2), mostramos que a sequência gerada pelo MPP é limitada.

Teorema 7. *Se as suposições (A1) e (A2) são válidas, então $\{\mathbf{d}(\mathbf{x}^k, \mathbf{x}^*)\}$ é convergente para cada $\mathbf{x}^* \in \mathbf{S}$. Em particular, $\{\mathbf{x}^k\}$ é limitada.*

Demonstração. De (2.6), temos que $\exp_{x^{k+1}}^{-1} \mathbf{y}^k \in \mu_k \mathcal{A}(x^{k+1})$ e consequentemente

$$\mathbf{y}^k \in \exp_{x^{k+1}} \mu_k \mathcal{A}(x^{k+1}).$$

Isso significa que $x^{k+1} = J_{\mu_k}^{\mathcal{A}}(\mathbf{y}^k)$ e por (2.5), temos

$$x^{k+1} = J_{\mu_k}^{\mathcal{A}}(\exp_{x^k} \mu_k \mathbf{w}^k),$$

onde $\mathbf{w}^k = \mathcal{B}(x^k)$. Seja $x^* \in S$ fixado e escolha $\mathbf{w}^* \in \mathcal{A}(x^*) \cap \mathcal{B}(x^*)$ tal que

$$x^* = J_{\mu_k}^{\mathcal{A}}(\exp_{x^*} \mu_k \mathbf{w}^*).$$

Sendo \mathcal{A} monótono maximal pela Proposição 10, seu resolvente é firmemente não-expansivo. Assim, usando o fato que o resolvente do campo \mathcal{A} é firmemente não expansivo, segue da Proposição 9 que

$$\begin{aligned} d(x^{k+1}, x^*) &= d(J_{\mu_k}^{\mathcal{A}}(\exp_{x^k} \mu_k \mathbf{w}^k), J_{\mu_k}^{\mathcal{A}}(\exp_{x^*} \mu_k \mathbf{w}^*)) \\ &\leq d(\exp_{x^k} \mu_k \mathbf{w}^k, \exp_{x^*} \mu_k \mathbf{w}^*). \end{aligned}$$

Aplicando agora, a desigualdade triangular no segundo membro da desigualdade acima, temos que

$$\begin{aligned} d(x^{k+1}, x^*) &\leq d(\exp_{x^k} \mu_k \mathbf{w}^k, x^*) + d(x^*, \exp_{x^*} \mu_k \mathbf{w}^*) \\ &\leq d(x^k, x^*) + d(\exp_{x^k} \mu_k \mathbf{w}^k, x^k) + d(x^*, \exp_{x^*} \mu_k \mathbf{w}^*). \end{aligned}$$

Segue da Observação 5 que $d(\exp_{x^k} \mu_k \mathbf{w}^k, x^k) = \mu_k \|\mathbf{w}^k\|$ e $d(x^*, \exp_{x^*} \mu_k \mathbf{w}^*) = \mu_k \|\mathbf{w}^*\|$, então aplicando novamente a desigualdade triangular e em seguida usando a condição (A1) e o fato que a aplicação transporte paralelo é uma isometria, obtemos

$$\begin{aligned} d(x^{k+1}, x^*) &\leq d(x^k, x^*) + \mu_k \|\mathbf{w}^k\| + \mu_k \|\mathbf{w}^*\| \\ &\leq d(x^k, x^*) + \mu_k (\|\mathbf{w}^k - \mathcal{P}_{x^k, x^*} \mathbf{w}^*\| + \|\mathcal{P}_{x^k, x^*} \mathbf{w}^*\|) + \mu_k \|\mathbf{w}^*\| \\ &\leq d(x^k, x^*) + \mu_k (Ld(x^k, x^*) + \|\mathbf{w}^*\|) + \mu_k \|\mathbf{w}^*\| \\ &= (1 + \mu_k L) d(x^k, x^*) + 2\mu_k \|\mathbf{w}^*\|. \end{aligned}$$

Portanto, aplicando o Lema 1 com $\alpha_k = d(x^k, x^*)$, $\beta_k = L\mu_k$ e $\gamma_k = 2\mu_k\|w^*\|$ levando em conta que (A2) é válida, então concluímos que $\{d(x^k, x^*)\}$ é convergente e, conseqüentemente $\{x^k\}$ é limitada. \square

O próximo resultado, mostra para a seqüência gerada pelo MPP, um comportamento clássico do método do ponto proximal, que é $\lim_{k \rightarrow +\infty} d(x^{k+1}, x^k) = 0$. Diferentemente da Proposição 13, o resultado não assume a limitação da seqüência ou a monotonicidade forte do campo vetorial B.

Corolário 1. *Se as suposições (A1) e (A2) são válidas, então $\lim_{k \rightarrow +\infty} d(x^{k+1}, x^k) = 0$.*

Demonstração. De (2.14), temos

$$\frac{1}{\mu_k} d^2(x^{k+1}, x^k) \leq \langle u^{k+1}, -\exp_{x^{k+1}}^{-1} x^k \rangle + \langle v^k, -\exp_{x^k}^{-1} x^{k+1} \rangle,$$

com $u^{k+1} = B(x^{k+1})$ e $v^k \in A(x^k)$.

Usando o transporte paralelo, temos que $\exp_{x^{k+1}}^{-1} x^k = -P_{x^{k+1}, x^k} \exp_{x^k}^{-1} x^{k+1}$ e então,

$$\frac{1}{\mu_k} d^2(x^{k+1}, x^k) \leq \|P_{x^{k+1}, x^k} v^k - u^{k+1}\| \|\exp_{x^{k+1}}^{-1} x^k\|.$$

Isso implica que

$$d(x^{k+1}, x^k) \leq \mu_k(\|v^k\| + \|u^{k+1}\|), \quad (2.22)$$

tendo em vista que $\|\exp_{x^{k+1}}^{-1} x^k\| = d(x^{k+1}, x^k) > 0$. Como, pelo Teorema 7, $\{x^k\}$ é limitada, temos pela Proposição 6 que as sequencias $\{u^k\}$ e $\{v^k\}$ são limitadas. Portanto, combinando (2.22) com (A2), temos que $\lim_{k \rightarrow +\infty} d(x^{k+1}, x^k) = 0$. \square

2.2.3 Condição suficiente para convergência total

Nesta subseção, apresentamos uma condição suficiente para obter a convergência total da seqüência gerada pelo nosso método. Sob as suposições (A1) e (A2), provamos pelo Corolário 1, que $\lim_{k \rightarrow +\infty} d(x^{k+1}, x^k) = 0$. Segue de (A2) que $\lim_{k \rightarrow +\infty} \mu_k = 0$. Sendo que esta seqüência auxiliar de parâmetros é escolhida livremente satisfazendo algumas condições adequadas, vamos tomá-la de tal forma que ela não vai a zero mais rápido que $d(x^{k+1}, x^k)$. Isso nos leva à seguinte condição:

(A3) Suponha que $\lim_{k \rightarrow \infty} \frac{d(x^{k+1}, x^k)}{\mu_k} = 0$.

Observação 17. *Uma suposição semelhante foi considerada, por exemplo, em (LEUŞTEAN et al., 2018) em espaços $CAT(0)$ (que inclui variedades de Hadamard). Em (LEUŞTEAN et al., 2018), assume-se que a sequência $\left\{ \frac{d(x^{k+1}, x^k)}{\mu_k} \right\}$ é não-crescente; veja (LEUŞTEAN et al., 2018, Condição (C2)). Além disso, se a sequência $\{\mu_k\}$ satisfaz $\sum_{k=0}^{\infty} \mu_k^2 = \infty$, eles obtêm que (A3) vale; veja (LEUŞTEAN et al., 2018, Lemma 3.3). Em (MOUDAFI, 2015) a condição (A3) também é assumida.*

Provamos abaixo que todo ponto de acumulação da sequência gerada pelo nosso método é um ponto crítico de $A - B$. Em seguida, mostramos que a sequência converge para um ponto crítico de $A - B$.

Teorema 8. *Se as suposições (A1) – (A3) são válidas, então todo ponto de acumulação de $\{x^k\}$ é ponto crítico de $A - B$.*

Demonstração. Seja $\bar{x} \in M$ um ponto de acumulação arbitrário de $\{x^k\}$ (o mesmo existe pelo Teorema 7), então existe $\{x^{k_j}\}$ subsequência de $\{x^k\}$ tal que $\lim_{j \rightarrow \infty} x^{k_j} = \bar{x}$. Pelo Corolário 1, temos que $\lim_{j \rightarrow \infty} x^{k_j+1} = \bar{x}$. Além disso, por (2.5) temos $w^{k_j} = B(x^{k_j})$, então pela Proposição 6, temos que $\{w^{k_j}\}$ é limitada. Portanto, sem perda de generalidade, podemos (se necessário) tomar uma subsequência $\{w^{k_j}\}$ convergindo para \bar{w} .

Sendo $w^{k_j} = B(x^{k_j})$, pela monotonicidade de B , temos

$$\langle w^{k_j}, \exp_{x^{k_j}}^{-1} z \rangle \leq \langle u, -\exp_z^{-1} x^{k_j} \rangle, \quad u = B(z), \quad \forall z \in M.$$

Fazendo $j \rightarrow \infty$, obtemos que

$$\langle \bar{w}, \exp_{\bar{x}}^{-1} z \rangle \leq \langle u, -\exp_z^{-1} \bar{x} \rangle, \quad u = B(z), \quad \forall z \in M.$$

Pela maximalidade de B , segue que $\bar{w} = B(\bar{x})$.

Por outro lado, de (2.6), temos

$$\frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} y^{k_j} \in A(x^{k_j+1}),$$

onde $y^{k_j} = \exp_{x^{k_j}} \mu_{k_j} w^{k_j}$. Afirmamos que $\lim_{j \rightarrow \infty} \frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} y^{k_j} = \bar{w}$, e assim, $\bar{w} \in A(\bar{x})$.

De fato,

$$\begin{aligned} \left\| \frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} - \mathbf{P}_{x^{k_j+1}, \bar{x}} \bar{\mathbf{w}} \right\| &\leq \left\| \frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} - \mathbf{P}_{x^{k_j+1}, x^{k_j}} \mathbf{w}^{k_j} \right\| \\ &\quad + \left\| \mathbf{P}_{x^{k_j+1}, x^{k_j}} \mathbf{w}^{k_j} - \mathbf{P}_{x^{k_j+1}, \bar{x}} \bar{\mathbf{w}} \right\|. \end{aligned} \quad (2.23)$$

Vamos provar que o lado direito da desigualdade acima se anula quando $j \rightarrow \infty$. Primeiro, note que pelo fato que a aplicação transporte paralelo é uma isometria e usando que $\mathbf{w}^{k_j} = \frac{1}{\mu_{k_j}} \exp_{x^{k_j}}^{-1} \mathbf{y}^{k_j}$, temos para o primeiro termo que

$$\begin{aligned} &\left\| \frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} - \mathbf{P}_{x^{k_j+1}, x^{k_j}} \mathbf{w}^{k_j} \right\|^2 = \\ &= \frac{1}{\mu_{k_j}^2} \left\| \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} \right\|^2 + \frac{1}{\mu_{k_j}^2} \left\| \exp_{x^{k_j}}^{-1} \mathbf{y}^{k_j} \right\|^2 - 2 \frac{1}{\mu_{k_j}} \left\langle \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j}, \mathbf{P}_{x^{k_j+1}, x^{k_j}} \mathbf{w}^{k_j} \right\rangle \\ &= \frac{1}{\mu_{k_j}^2} \left(d^2(x^{k_j+1}, \mathbf{y}^{k_j}) + d^2(\mathbf{y}^{k_j}, x^{k_j}) \right) - 2 \frac{1}{\mu_{k_j}} \left\langle \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j}, \mathbf{P}_{x^{k_j+1}, x^{k_j}} \mathbf{w}^{k_j} \right\rangle. \end{aligned}$$

Usando mais uma vez a isometria do transporte paralelo e as seguintes igualdades

$$\mathbf{P}_{\mathbf{y}^{k_j}, x^{k_j+1}} \left(\mathbf{P}_{x^{k_j+1}, x^{k_j}} \right) = \mathbf{P}_{\mathbf{y}^{k_j}, x^{k_j}} \text{ e } \mathbf{P}_{\mathbf{y}^{k_j}, x^{k_j+1}} \left(\exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} \right) = -\exp_{\mathbf{y}^{k_j}}^{-1} x^{k_j+1}, \text{ segue que}$$

$$\begin{aligned} &\left\| \frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} - \mathbf{P}_{x^{k_j+1}, x^{k_j}} \mathbf{w}^{k_j} \right\|^2 = \\ &= \frac{1}{\mu_{k_j}^2} \left(d^2(x^{k_j+1}, \mathbf{y}^{k_j}) + d^2(\mathbf{y}^{k_j}, x^{k_j}) \right) \\ &\quad - \frac{2}{\mu_{k_j}} \left\langle \mathbf{P}_{\mathbf{y}^{k_j}, x^{k_j+1}} \left(\exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} \right), \mathbf{P}_{\mathbf{y}^{k_j}, x^{k_j+1}} \left(\mathbf{P}_{x^{k_j+1}, x^{k_j}} \mathbf{w}^{k_j} \right) \right\rangle \\ &= \frac{1}{\mu_{k_j}^2} \left(d^2(x^{k_j+1}, \mathbf{y}^{k_j}) + d^2(\mathbf{y}^{k_j}, x^{k_j}) \right) - \frac{2}{\mu_{k_j}} \left\langle -\exp_{\mathbf{y}^{k_j}}^{-1} x^{k_j+1}, \mathbf{P}_{\mathbf{y}^{k_j}, x^{k_j}} \left(\frac{1}{\mu_{k_j}} \exp_{x^{k_j}}^{-1} \mathbf{y}^{k_j} \right) \right\rangle. \end{aligned}$$

Sendo o transporte paralelo uma isometria linear, obtemos que

$$\mathbf{P}_{\mathbf{y}^{k_j}, x^{k_j}} \left(\frac{1}{\mu_{k_j}} \exp_{x^{k_j}}^{-1} \mathbf{y}^{k_j} \right) = \frac{1}{\mu_{k_j}} \mathbf{P}_{\mathbf{y}^{k_j}, x^{k_j}} \left(\exp_{x^{k_j}}^{-1} \mathbf{y}^{k_j} \right) = -\frac{1}{\mu_{k_j}} \exp_{\mathbf{y}^{k_j}}^{-1} x^{k_j},$$

implicando que

$$\begin{aligned}
 & \left\| \frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} - \mathbf{P}_{x^{k_j+1}, x^{k_j}} \mathbf{w}^{k_j} \right\|^2 = \\
 &= \frac{1}{\mu_{k_j}^2} \left(d^2(x^{k_j+1}, \mathbf{y}^{k_j}) + d^2(\mathbf{y}^{k_j}, x^{k_j}) \right) - 2 \frac{1}{\mu_{k_j}} \left\langle -\exp_{y^{k_j}}^{-1} x^{k_j+1}, -\frac{1}{\mu_{k_j}} \exp_{y^{k_j}}^{-1} x^{k_j} \right\rangle \\
 &= \frac{1}{\mu_{k_j}^2} \left(d^2(x^{k_j+1}, \mathbf{y}^{k_j}) + d^2(\mathbf{y}^{k_j}, x^{k_j}) \right) - 2 \frac{1}{\mu_{k_j}^2} \left\langle \exp_{y^{k_j}}^{-1} x^{k_j+1}, \exp_{y^{k_j}}^{-1} x^{k_j} \right\rangle. \quad (2.24)
 \end{aligned}$$

Do triângulo geodésico $\triangle(x^{k_j+1}, \mathbf{y}^{k_j}, x^{k_j})$, temos

$$d^2(x^{k_j+1}, \mathbf{y}^{k_j}) + d^2(\mathbf{y}^{k_j}, x^{k_j}) - 2 \langle \exp_{y^{k_j}}^{-1} x^{k_j+1}, \exp_{y^{k_j}}^{-1} x^{k_j} \rangle \leq d^2(x^{k_j+1}, x^{k_j}).$$

Usando a última desigualdade em (2.24), temos

$$\left\| \frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} - \mathbf{P}_{x^{k_j+1}, x^{k_j}} \mathbf{w}^{k_j} \right\| \leq \frac{1}{\mu_{k_j}} d(x^{k_j+1}, x^{k_j}).$$

Fazendo $j \rightarrow \infty$ na desigualdade acima e usando (A3), segue que

$$\lim_{j \rightarrow \infty} \left\| \frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} - \mathbf{P}_{x^{k_j+1}, x^{k_j}} \mathbf{w}^{k_j} \right\| = 0. \quad (2.25)$$

Agora, observe que

$$\| \mathbf{P}_{x^{k_j+1}, x^{k_j}} \mathbf{w}^{k_j} - \mathbf{P}_{x^{k_j+1}, \bar{x}} \bar{\mathbf{w}} \| = \| \mathbf{P}_{x^{k_j+1}, x^{k_j}} \mathbf{w}^{k_j} - \mathbf{P}_{x^{k_j+1}, x^{k_j}} (\mathbf{P}_{x^{k_j}, \bar{x}} \bar{\mathbf{w}}) \| = \| \mathbf{w}^{k_j} - \mathbf{P}_{x^{k_j}, \bar{x}} \bar{\mathbf{w}} \|.$$

Como $x^{k_j} \rightarrow \bar{x}$ e $\mathbf{w}^{k_j} \rightarrow \bar{\mathbf{w}}$ fazendo $j \rightarrow \infty$ na igualdade acima obtemos

$$\lim_{j \rightarrow \infty} \| \mathbf{P}_{x^{k_j+1}, x^{k_j}} \mathbf{w}^{k_j} - \mathbf{P}_{x^{k_j+1}, \bar{x}} \bar{\mathbf{w}} \| = 0. \quad (2.26)$$

Portanto, usando (2.25) e (2.26) em (2.23), temos

$$\lim_{j \rightarrow \infty} \frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} = \bar{\mathbf{w}}.$$

Usando agora a monotonicidade de \mathbf{A} , temos que

$$\left\langle \frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j}, \exp_{x^{k_j+1}}^{-1} z \right\rangle \leq \langle \mathbf{u}, -\exp_z^{-1} x^{k_j+1} \rangle, \quad \forall \mathbf{u} \in \mathbf{A}(z), z \in M.$$

Fazendo $j \rightarrow \infty$, obtemos

$$\langle \bar{w}, \exp_{\bar{x}}^{-1} z \rangle \leq \langle u, -\exp_z^{-1} \bar{x} \rangle, \quad \forall u \in A(z), z \in M.$$

Pela maximalidade de A , segue que $\bar{w} \in A(\bar{x})$ e assim $\bar{w} \in A(\bar{x}) \cap B(\bar{x})$ implicando que \bar{x} é um ponto crítico de $A - B$, isto é, $\bar{x} \in S$. \square

Mostraremos agora a convergência total da sequência para um ponto crítico de $A - B$.

Teorema 9. *Se as suposições (A1)–(A3) são válidas, então $\{x^k\}$ converge para um ponto crítico de $A - B$.*

Demonstração. De fato, em virtude do Teorema 7 a sequência $\{x^k\}$ é limitada. Seja $\{x^{k_j}\}$ subsequência de $\{x^k\}$ tal que $\lim_{j \rightarrow \infty} x^{k_j} = x^* \in M$. Assim,

$$\lim_{j \rightarrow \infty} d(x^{k_j}, x^*) = 0.$$

Como x^* é um ponto de acumulação de $\{x^k\}$, segue do Teorema 8 que $x^* \in S$. Então, pelo Teorema 7, a sequência $\{d(x^k, x^*)\}$ é convergente implicando que

$$\lim_{k \rightarrow \infty} d(x^k, x^*) = 0.$$

Portanto, $\{x^k\}$ converge para $x^* \in S$, ou seja, $\{x^k\}$ converge para um ponto crítico de $A - B$. \square

Capítulo 3

Método do ponto proximal inercial

Neste capítulo, apresentamos um método de ponto proximal inercial (MPPI) para resolver o Problema (2.3), ou seja,

$$\text{encontrar } \mathbf{x}^* \in M \text{ tal que } A(\mathbf{x}^*) \cap B(\mathbf{x}^*) \neq \emptyset,$$

via generalização do MPP, estudado no capítulo anterior. Os resultados deste capítulo encontram-se em (ANDRADE et al., 2022).

No cenário linear, (ALIMOHAMMADY; RAMAZANNEJAD, 2016) propuseram um algoritmo de ponto proximal inercial para a diferença de operadores monótonos maximais baseado no método considerado por (MOUDAFI, 2015). Esses métodos são baseados no operador resolvente e na aproximação de Yosida. Vale ressaltar que no cenário linear nossos métodos são diferentes dos algoritmos propostos em (MOUDAFI, 2008, 2015; ALIMOHAMMADY; RAMAZANNEJAD, 2016).

3.1 Algoritmo MPPI

Lembrando que M é uma variedade de Hadamard de dimensão finita, A e $B \in \mathcal{X}(M)$ são campos vetoriais monótonos maximais com $D(A) = D(B) = M$, o conjunto solução de (2.3), isto é,

$$S = \{\mathbf{x}^* \in M; A(\mathbf{x}^*) \cap B(\mathbf{x}^*) \neq \emptyset\},$$

o conjunto dos pontos críticos de $A - B$ é considerado não vazio.

Baseado no MPP, temos o seguinte método do ponto proximal inercial:

Método do ponto proximal inercial - MPPI

Passo 1: Dados um ponto inicial $\mathbf{x}^0 \in M$, $\rho > 0$, $\gamma_k \in [0, \frac{\rho}{2})$ e considere $\mathbf{a}, \mathbf{b} > 0$. Defina $\mathbf{x}^{-1} = \mathbf{x}^0$ e seja $\{\mu_k\}$ uma sequência de números positivos tal que

$$\mathbf{a} \leq \mu_k \leq \mathbf{b}. \quad (3.1)$$

Passo 2: Dado $\mathbf{x}^k \in M$, defina $\mathbf{d}^k = \gamma_k \exp_{\mathbf{x}^k}^{-1} \mathbf{x}^{k-1}$. Escolha

$$\mathbf{w}^k \in B(\mathbf{x}^k) \text{ e defina } \mathbf{y}^k := \exp_{\mathbf{x}^k} \mu_k (\mathbf{w}^k + \mathbf{d}^k). \quad (3.2)$$

Passo 3: Calcule $\mathbf{x}^{k+1} \in M$ tal que

$$0 \in A(\mathbf{x}^{k+1}) - \frac{1}{\mu_k} \exp_{\mathbf{x}^{k+1}}^{-1} \mathbf{y}^k. \quad (3.3)$$

Se $\mathbf{x}^{k+1} = \mathbf{x}^k$ e $\mathbf{d}^k = 0$, pare. Caso contrário, defina $k := k + 1$ e retorne ao Passo 2.

Observação 18. Observe que se $\gamma_k = 0$, para todo $k \in \mathbb{N}$ no MPPI, então $\mathbf{d}^k = 0$ e (3.2) se torna

$$\mathbf{w}^k \in B(\mathbf{x}^k) \text{ e } \mathbf{y}^k = \exp_{\mathbf{x}^k} \mu_k \mathbf{w}^k.$$

Neste caso, o MPPI reduz-se ao MPP. Sendo, $\mathbf{x}^{k+1} \in M$ definido como (3.3), se $B(\mathbf{x}) = \partial h(\mathbf{x})$ e $A(\mathbf{x}) = \partial g(\mathbf{x})$, onde $g, h : M \rightarrow \mathbb{R}$ são funções convexas, então (3.3) é equivalente a

$$\mathbf{x}^{k+1} = \arg \min_{\mathbf{x} \in M} \{g(\mathbf{x}) + \frac{1}{\mu_k} d^2(\mathbf{x}, \mathbf{y}^k)\}.$$

Portanto, vemos que neste caso, o MPPI torna-se o MPP para funções DC proposto por (SOUZA; OLIVEIRA, 2015). Assim, MPPI com $\gamma_k = 0$, para todo $k \in \mathbb{N}$, é uma generalização do MPP e pode ser visto como uma extensão à diferença de campos vetoriais monótonos maximais do método considerado em (SOUZA; OLIVEIRA, 2015) para diferença de funções convexas. Neste caso, nosso método é diferente dos métodos regularizados para a diferença de operadores monótonos máximos propostos por Moudafi (MOUDAFI, 2008, 2015) na configuração linear.

Proposição 14. As sequências $\{\mathbf{y}^k\}$ e $\{\mathbf{x}^k\}$ geradas pelo MPPI são bem definidas. Além disso, se $\mathbf{x}^{k+1} = \mathbf{x}^k$ e $\mathbf{d}^k = 0$ o MPPI para em um solução de (2.3).

Demonstração. A boa definição das sequências $\{\mathbf{x}^k\}$ e $\{\mathbf{y}^k\}$ segue da Proposição 11. Para a prova da segunda parte, temos por (3.2) e (3.3) que

$$\frac{1}{\mu_k} \exp_{\mathbf{x}^k}^{-1} \mathbf{y}^k - \mathbf{d}^k \in B(\mathbf{x}^k) \quad \text{e} \quad \frac{1}{\mu_k} \exp_{\mathbf{x}^{k+1}}^{-1} \mathbf{y}^k \in A(\mathbf{x}^{k+1}).$$

Assim, se $\mathbf{x}^{k+1} = \mathbf{x}^k$ e $\mathbf{d}^k = 0$, obtemos

$$\frac{1}{\mu_k} \exp_{\mathbf{x}^k}^{-1} \mathbf{y}^k \in B(\mathbf{x}^k) \quad \text{e} \quad \frac{1}{\mu_k} \exp_{\mathbf{x}^k}^{-1} \mathbf{y}^k \in A(\mathbf{x}^k).$$

Portanto, $\frac{1}{\mu_k} \exp_{\mathbf{x}^k}^{-1} \mathbf{y}^k \in A(\mathbf{x}^k) \cap B(\mathbf{x}^k)$, isto é, $A(\mathbf{x}^k) \cap B(\mathbf{x}^k) \neq \emptyset$ e conseqüentemente \mathbf{x}^k é uma solução de (2.3). \square

Agora, apresentamos uma análise de convergência para o MPPI para a diferença de campos vetoriais monótonos maximais (que é um campo vetorial possivelmente não monótono).

3.2 Análise de convergência: caso possivelmente não monótono

O objetivo desta seção é estudar as propriedades de convergência do MPPI. A partir de agora, consideramos $\{\mathbf{x}^k\}$ a sequência gerada pelo MPPI e assumimos que $\mathbf{x}^{k+1} \neq \mathbf{x}^k$, para todo $k \in \mathbb{N}$. Agora vamos estabelecer suas propriedades de convergência. Semelhante ao MPP, provamos a seguinte proposição que desempenha um papel importante na análise de convergência do MPPI.

Proposição 15. *Suponha que A ou B é ρ -fortemente monótono. Então,*

$$\left(\rho + \frac{1}{b} \right) d(\mathbf{x}^{k+1}, \mathbf{x}^k) \leq \|P_{\mathbf{x}^k, \mathbf{x}^{k+1}} \mathbf{u}^{k+1} - \mathbf{v}^k\| + \frac{\rho}{2} d(\mathbf{x}^k, \mathbf{x}^{k-1}), \quad (3.4)$$

para quaisquer $\mathbf{v}^k \in A(\mathbf{x}^k)$ e $\mathbf{u}^{k+1} \in B(\mathbf{x}^{k+1})$.

Demonstração. Vamos supor que A é monótono e B é ρ -fortemente monótono. O outro caso é análogo. Por (3.2) e (3.3), temos

$$\frac{1}{\mu_k} \exp_{\mathbf{x}^k}^{-1} \mathbf{y}^k - \mathbf{d}^k \in B(\mathbf{x}^k) \quad \text{e} \quad \frac{1}{\mu_k} \exp_{\mathbf{x}^{k+1}}^{-1} \mathbf{y}^k \in A(\mathbf{x}^{k+1}).$$

Dado $x \in M$, como B é ρ -fortemente monótono, então

$$\left\langle \frac{1}{\mu_k} \exp_{x^k}^{-1} y^k - d^k, \exp_{x^k}^{-1} x \right\rangle - \langle u, -\exp_{x^k}^{-1} x^k \rangle \leq -\rho d^2(x, x^k), \quad \forall u \in B(x). \quad (3.5)$$

Segue da monotonicidade de A que

$$\left\langle \frac{1}{\mu_k} \exp_{x^{k+1}}^{-1} y^k, \exp_{x^{k+1}}^{-1} x \right\rangle \leq \langle v, -\exp_{x^{k+1}}^{-1} x^{k+1} \rangle, \quad \forall v \in A(x). \quad (3.6)$$

Tomando $x = x^{k+1}$ em (3.5), $x = x^k$ em (3.6) e somando essas desigualdades, temos

$$\frac{1}{\mu_k} [\langle \exp_{x^k}^{-1} y^k - d^k, \exp_{x^k}^{-1} x^{k+1} \rangle + \langle \exp_{x^{k+1}}^{-1} y^k, \exp_{x^{k+1}}^{-1} x^k \rangle] \leq \quad (3.7)$$

$$\langle u^{k+1}, -\exp_{x^{k+1}}^{-1} x^k \rangle + \langle v^k, -\exp_{x^k}^{-1} x^{k+1} \rangle - \rho d^2(x^{k+1}, x^k) + \langle d^k, \exp_{x^k}^{-1} x^{k+1} \rangle,$$

para quaisquer $v^k \in A(x^k)$ e $u^{k+1} \in B(x^{k+1})$. Seja $\Delta(y^k, x^k, x^{k+1})$ o triângulo geodésico com $\theta = \angle(\exp_{x^k}^{-1} y^k, \exp_{x^k}^{-1} x^{k+1})$. Então, pelo Teorema 3 (Teorema de comparação para triângulos), temos

$$d^2(y^k, x^k) + d^2(x^k, x^{k+1}) - 2\langle \exp_{x^k}^{-1} y^k, \exp_{x^k}^{-1} x^{k+1} \rangle \leq d^2(y^k, x^{k+1}). \quad (3.8)$$

Similarmente, para o triângulo geodésico $\Delta(y^k, x^{k+1}, x^k)$ com $\theta = \angle(\exp_{x^{k+1}}^{-1} y^k, \exp_{x^{k+1}}^{-1} x^k)$, temos

$$d^2(y^k, x^{k+1}) + d^2(x^{k+1}, x^k) - 2\langle \exp_{x^{k+1}}^{-1} y^k, \exp_{x^{k+1}}^{-1} x^k \rangle \leq d^2(y^k, x^k). \quad (3.9)$$

Adicionando (3.8) e (3.9), obtemos

$$d^2(x^{k+1}, x^k) \leq \langle \exp_{x^k}^{-1} y^k, \exp_{x^k}^{-1} x^{k+1} \rangle + \langle \exp_{x^{k+1}}^{-1} y^k, \exp_{x^{k+1}}^{-1} x^k \rangle. \quad (3.10)$$

Sendo $\mu_k > 0$, para todo $k \in \mathbb{N}$, e usando (3.10) em (3.7), temos

$$\frac{1}{\mu_k} d^2(x^{k+1}, x^k) \leq$$

$$\langle u^{k+1}, -\exp_{x^{k+1}}^{-1} x^k \rangle + \langle v^k, -\exp_{x^k}^{-1} x^{k+1} \rangle - \rho d^2(x^{k+1}, x^k) + \langle d^k, \exp_{x^k}^{-1} x^{k+1} \rangle. \quad (3.11)$$

Assim, usando o fato que $0 < \mu_k \leq b$, obtemos

$$\left(\rho + \frac{1}{b}\right) d^2(x^{k+1}, x^k) \leq \langle u^{k+1}, -\exp_{x^{k+1}}^{-1} x^k \rangle + \langle v^k, -\exp_{x^k}^{-1} x^{k+1} \rangle + \langle d^k, \exp_{x^k}^{-1} x^{k+1} \rangle, \quad (3.12)$$

para quaisquer $v^k \in A(x^k)$ e $u^{k+1} \in B(x^{k+1})$.

Por outro lado, usando as propriedades de transporte paralelo, temos que $\exp_{x^k}^{-1} x^{k+1} = -P_{x^k, x^{k+1}} \exp_{x^{k+1}}^{-1} x^k$ e, conseqüentemente,

$$\langle u^{k+1}, -\exp_{x^{k+1}}^{-1} x^k \rangle + \langle v^k, -\exp_{x^k}^{-1} x^{k+1} \rangle = \langle P_{x^k, x^{k+1}} u^{k+1} - v^k, \exp_{x^k}^{-1} x^{k+1} \rangle.$$

Portanto, usando este fato em (3.12), tem-se

$$\left(\rho + \frac{1}{b}\right) d^2(x^{k+1}, x^k) \leq \langle P_{x^k, x^{k+1}} u^{k+1} - v^k + d^k, \exp_{x^k}^{-1} x^{k+1} \rangle, \quad (3.13)$$

para quaisquer $v^k \in A(x^k)$ e $u^{k+1} \in B(x^{k+1})$.

Aplicando agora a desigualdade de Cauchy-Schwarz e a desigualdade triangular em (3.13), temos

$$\left(\rho + \frac{1}{b}\right) d^2(x^{k+1}, x^k) \leq \left(\|P_{x^k, x^{k+1}} u^{k+1} - v^k\| + \|d^k\| \right) \| \exp_{x^k}^{-1} x^{k+1} \|,$$

para quaisquer $v^k \in A(x^k)$ e $u^{k+1} \in B(x^{k+1})$.

Usando o fato que, em variedades de Hadamard, temos $\| \exp_{x^k}^{-1} x^{k+1} \| = d(x^{k+1}, x^k)$ e $\|d^k\| = \|\gamma_k \exp_{x^k}^{-1} x^{k-1}\| = \gamma_k d(x^k, x^{k-1})$, segue que

$$\left(\rho + \frac{1}{b}\right) d(x^{k+1}, x^k) \leq \|P_{x^k, x^{k+1}} u^{k+1} - v^k\| + \gamma_k d(x^k, x^{k-1}),$$

para quaisquer $v^k \in A(x^k)$ e $u^{k+1} \in B(x^{k+1})$. Portanto, o resultado segue usando o fato que $\gamma_k \in [0, \frac{\rho}{2})$. \square

Observação 19. *Relembramos que, embora em nossa prova precisamos da monotonicidade forte de apenas um dos campos vetoriais A e B , vale a pena mencionar que a monotonicidade forte de A e B não é uma suposição restritiva, veja Observação 13.*

Mostramos também, que a distância Riemanniana entre duas iteradas consecutivas do MPPI tende para zero quando k tende para ∞ , que é uma propriedade clássica de

métodos de pontos proximal.

Proposição 16. *Suponha que $\{x^k\}$ é limitada, então existem constantes $L, M \geq 0$ tais que $\|P_{x^k, x^{k+1}}u^{k+1} - v^k\| \leq L$ e $\limsup_{k \rightarrow \infty} d(x^{k+1}, x^k) = M$, para quaisquer $v^k \in A(x^k)$, $u^{k+1} \in B(x^{k+1})$ para todo k . Além disso, se A ou B é ρ -fortemente monótono e*

$$\left(\frac{\rho}{2} + \frac{1}{b}\right)M > L, \quad (3.14)$$

então $\lim_{k \rightarrow \infty} d(x^{k+1}, x^k) = 0$. Consequentemente $\lim_{k \rightarrow \infty} \|d^k\| = 0$.

Demonstração. Como $\{x^k\}$ é limitada segue da maximalidade de A e B (veja Proposição 6) que existem constantes $K_1, K_2 > 0$ tais que $\|v^k\| < K_1$ e $\|u^k\| < K_2$ para todo k . Assim, pela desigualdade triangular, temos

$$\|P_{x^k, x^{k+1}}u^{k+1} - v^k\| \leq \|P_{x^k, x^{k+1}}u^{k+1}\| + \|v^k\|.$$

Usando o fato que a aplicação transporte paralelo é uma isometria, temos

$$\|P_{x^k, x^{k+1}}u^{k+1} - v^k\| < L := K_1 + K_2, \quad (3.15)$$

para todo k , e conseqüentemente, a desigualdade está provada. Por outro lado, sendo $\{x^k\}$ limitada, então a sequência $\{d(x^{k+1}, x^k)\}$ é limitada. Denote por $M = \limsup_{k \rightarrow \infty} d(x^{k+1}, x^k)$. Assuma por contradição que $M > 0$ e, suponha que A ou B é ρ -fortemente monótono com $\rho > 0$ satisfazendo (3.14). Assim, combinando (3.4) e (3.15), temos

$$\left(\rho + \frac{1}{b}\right)d(x^{k+1}, x^k) \leq \|P_{x^{k+1}, x^k}v^k - u^{k+1}\| + \frac{\rho}{2}d(x^k, x^{k-1}) \leq L + \frac{\rho}{2}d(x^k, x^{k-1}).$$

Tomando o $\limsup_{k \rightarrow \infty}$ na desigualdade acima, isso implica que $\left(\frac{\rho}{2} + \frac{1}{b}\right)M \leq L$, o que é uma contradição. Portanto, a afirmação está provada. Portanto, $\lim_{k \rightarrow \infty} d(x^{k+1}, x^k) = 0$. Consequentemente $\lim_{k \rightarrow \infty} \|d^k\| = \lim_{k \rightarrow \infty} \|\gamma_k \exp_{x^k}^{-1} x^{k-1}\| = \lim_{k \rightarrow \infty} \gamma_k d(x^k, x^{k-1}) = 0$, onde foi usado o fato que $d(x^k, x^{k-1}) = \|\exp_{x^k}^{-1} x^{k-1}\|$ e $\gamma_k \in [0, \frac{\rho}{2}]$. \square

Observação 20. *Semelhante ao que foi comentado na Observação 14, se $A(\cdot) = \partial g(\cdot)$ e $B(\cdot) = \partial h(\cdot)$ com g e h funções convexas e, conseqüentemente A e B são campos vetoriais monótonos maximais. Sendo $f(x) = g(x) - h(x)$, então (3.14) é substituída pela*

limitação inferior de f . No entanto, a prova ainda segue de uma desigualdade como (3.4), mais precisamente, $\frac{1}{\mathbf{b}} \mathbf{d}^2(\mathbf{x}^{k+1}, \mathbf{x}^k) \leq f(\mathbf{x}^0) - f^*$, onde $f^* = \inf_{\mathbf{x} \in \mathbf{M}} f(\mathbf{x})$; veja (SOUZA; OLIVEIRA, 2015, Proposição 5). Lidando com campos vetoriais monótonos maximais gerais, vale a pena mencionar que a condição (3.14) não é restritiva porque o parâmetro ρ pode ser tomado grande o suficiente para que (3.14) seja válida; veja Observação 13.

3.2.1 Análise de convergência parcial

Nesta subseção, apresentamos resultados de convergência parcial para o MPPI. Na subseção seguinte, apresentamos resultados de convergência total para MPPI para o caso mais geral de diferença de campos vetoriais monótonos em variedades de Hadamard. Além disso, introduzimos uma suposição sobre o parâmetro proximal $\{\mu_k\}$, sobre a sequência $\{\gamma_k \mathbf{d}(\mathbf{x}^k, \mathbf{x}^{k-1})\}$ e uma continuidade Lipschitz de \mathbf{B} para obter convergência total da sequência para um ponto crítico de $\mathbf{A} - \mathbf{B}$.

Teorema 10. *Suponha que $\{\mathbf{x}^k\}$ é limitada e \mathbf{A} ou \mathbf{B} é ρ -fortemente monótono. Então, todo ponto de acumulação de $\{\mathbf{x}^k\}$ é um ponto crítico de $\mathbf{A} - \mathbf{B}$.*

Demonstração. Considere \mathbf{x} , um ponto de acumulação de $\{\mathbf{x}^k\}$, e seja $\{\mathbf{x}^{k_j}\}$ subsequência de $\{\mathbf{x}^k\}$ convergindo para \mathbf{x} . Sendo $\{\mu_k\}$ e $\{\mathbf{y}^k\}$ limitadas, podemos tomar sem perda de generalidade subsequências $\{\mu_{k_j}\}$ e $\{\mathbf{y}^{k_j}\}$ convergindo respectivamente para μ e \mathbf{y} . Pela Proposição 16, temos que $\mathbf{x}^{k_j+1} \rightarrow \mathbf{x}$. Segue por (3.2) que

$$\frac{1}{\mu_{k_j}} \exp_{\mathbf{x}^{k_j}}^{-1} \mathbf{y}^{k_j} - \mathbf{d}^{k_j} \in \mathbf{B}(\mathbf{x}^{k_j}).$$

Pela monotonicidade de \mathbf{B} , temos

$$\left\langle \frac{1}{\mu_{k_j}} \exp_{\mathbf{x}^{k_j}}^{-1} \mathbf{y}^{k_j} - \mathbf{d}^{k_j}, \exp_{\mathbf{x}^{k_j}}^{-1} \mathbf{z} \right\rangle \leq \langle \mathbf{u}, -\exp_{\mathbf{z}}^{-1} \mathbf{x}^{k_j} \rangle, \quad \forall \mathbf{u} \in \mathbf{B}(\mathbf{z}), \mathbf{z} \in \mathbf{M}.$$

Fazendo $j \rightarrow +\infty$ na última desigualdade e usando a Proposição 1, obtemos

$$\left\langle \frac{1}{\mu} \exp_{\mathbf{x}}^{-1} \mathbf{y}, \exp_{\mathbf{x}}^{-1} \mathbf{z} \right\rangle \leq \langle \mathbf{u}, -\exp_{\mathbf{z}}^{-1} \mathbf{x} \rangle, \quad \forall \mathbf{u} \in \mathbf{B}(\mathbf{z}), \mathbf{z} \in \mathbf{M},$$

e conseqüentemente, pela maximalidade de \mathbf{B} , temos

$$\frac{1}{\mu} \exp_{\mathbf{x}}^{-1} \mathbf{y} \in \mathbf{B}(\mathbf{x}). \tag{3.16}$$

Agora, de (3.3), temos

$$\frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} \in \mathcal{A}(x^{k_j+1}).$$

Similarmente, pela monotonicidade de \mathcal{A} , temos

$$\left\langle \frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j}, \exp_{x^{k_j+1}}^{-1} z \right\rangle \leq \langle \mathbf{u}, -\exp_z^{-1} x^{k_j+1} \rangle, \quad \forall \mathbf{u} \in \mathcal{A}(z), \forall z \in M,$$

e fazendo $j \rightarrow +\infty$, obtemos que

$$\left\langle \frac{1}{\mu} \exp_x^{-1} \mathbf{y}, \exp_x^{-1} z \right\rangle \leq \langle \mathbf{u}, -\exp_z^{-1} x \rangle, \quad \forall \mathbf{u} \in \mathcal{A}(z), \forall z \in M.$$

Segue da maximalidade de \mathcal{A} que

$$\frac{1}{\mu} \exp_x^{-1} \mathbf{y} \in \mathcal{A}(x). \quad (3.17)$$

Portanto, de (3.16) e (3.17), temos $\frac{1}{\mu} \exp_x^{-1} \mathbf{y} \in \mathcal{A}(x) \cap \mathcal{B}(x)$ o que significa que (2.3) é satisfeito, isto é, x é ponto crítico de $\mathcal{A} - \mathcal{B}$. \square

Observação 21. *Assim como no Teorema 6 a suposição de ρ -forte monotonicidade de \mathcal{A} ou \mathcal{B} foi usada na prova do Teorema 10 apenas para garantir que se $x^{k_j} \rightarrow x$ com $j \rightarrow \infty$, então $x^{k_j+1} \rightarrow x$. Em outras palavras, a hipótese acima foi assumido para aplicar a Proposição 16, ou seja, para obter que $\lim_{k \rightarrow \infty} d(x^{k+1}, x^k) = 0$; veja Observação 15.*

3.2.2 Condições suficientes para a limitação

Como mencionado no Capítulo 2, recentemente, alguns trabalhos comprovaram a convergência de toda a sequência para diversos métodos aplicados a funções DC supondo a propriedade Kurdyka-Łojasiewicz da função objetivo; veja (ARTACHO; VUONG, 2020; CRUZ NETO et al., 2020; FERREIRA et al., 2021; THI et al., 2018). No entanto, mesmo neste caso, supõe-se que a sequência gerada pelo método seja limitada. O objetivo desta subseção é introduzir condições suficientes para garantir a limitação da sequência gerada pelo nosso método. Vale ressaltar que este resultado é novo mesmo para problemas DC no cenário Euclidiano. Para isso, vamos considerar nesta seção o MPPI substituindo a

condição (3.1) pela condição mais fraca:

$$0 < \mu_k \leq b. \quad (3.18)$$

Consideramos também as seguintes suposições sobre o campo vetorial B , os parâmetros proximal μ_k e a sequência de números não negativos $\{\gamma_k d(x^k, x^{k-1})\}$.

(A1) Dados $x, y \in M$, existe uma constante $L > 0$ tal que $\|u - P_{x,y}v\| \leq Ld(x, y)$, para quaisquer $u \in B(x)$ e $v \in B(y)$.

(A2) A sequência de números positivos $\{\mu_k\}$ satisfaz $\sum_{k=0}^{\infty} \mu_k < \infty$.

(A3) A sequência $\{\gamma_k d(x^k, x^{k-1})\}$ é limitada.

Observação 22. *Como mencionado na Observação 16, (A1) é uma suposição do tipo Lipschitz para campos vetoriais, ou seja, a suposição (A1) é uma extensão natural da continuidade de Lipschitz do gradiente da função DC para a diferença de campos vetoriais monótonos maximais. No contexto de campos vetoriais, a condição (A1) implica que o campo vetorial B é ponto-ponto. Já, a condição (A2) é uma suposição clássica em métodos de pontos proximal; veja por exemplo (FERREIRA; OLIVEIRA, 2002). Lidando com funções DC tem sido natural supor a limitação da sequência para diferentes métodos; veja por exemplo (SOUZA; OLIVEIRA, 2015; MAINGÉ; MOUDAFI, 2008; OLIVEIRA; TCHEOU, 2019; ARTACHO; VUONG, 2020; CRUZ NETO et al., 2020; FERREIRA et al., 2021; THI et al., 2018; ALMEIDA et al., 2020; BENTO et al., 2015). Alguns trabalhos substituíram essa suposição pelo conceito de desigualdade Kurdyka-Lojasiewicz; veja por exemplo (ARTACHO; VUONG, 2020; CRUZ NETO et al., 2020; FERREIRA et al., 2021; THI et al., 2018). Aqui, consideramos a suposição de limitação (A3) para obter que $\{x^k\}$ é limitada. É bastante natural porque $\gamma_k \in [0, \frac{\rho}{2})$, para todo $k \in \mathbb{N}$, e o método do ponto proximal e suas variantes geralmente têm a propriedade que $\{d(x^k, x^{k-1})\}$ converge para zero.*

Teorema 11. *Suponha que $S \neq \emptyset$. Se as suposições (A1), (A2) e (A3) são válidas, então $\{d(x^k, x^*)\}$ é convergente para cada $x^* \in S$. Em particular, $\{x^k\}$ é limitada.*

Demonstração. Por (3.3), temos que $\exp_{x^{k+1}}^{-1} y^k \in \mu_k A(x^{k+1})$ e conseqüentemente

$$y^k \in \exp_{x^{k+1}} \mu_k A(x^{k+1}).$$

Isso significa que $\mathbf{x}^{k+1} = J_{\mu_k}^A(\mathbf{y}^k)$ e de (3.2), temos

$$\mathbf{x}^{k+1} = J_{\mu_k}^A(\exp_{\mathbf{x}^k} \mu_k(\mathbf{w}^k + \mathbf{d}^k)),$$

onde $\mathbf{w}^k = B(\mathbf{x}^k)$. Seja $\mathbf{x}^* \in S$ fixado e escolha $\mathbf{w}^* \in A(\mathbf{x}^*) \cap B(\mathbf{x}^*)$ tal que

$$\mathbf{x}^* = J_{\mu_k}^A(\exp_{\mathbf{x}^*} \mu_k \mathbf{w}^*).$$

Como A é monótono maximal, pela Proposição 10, seu resolvente é firmemente não expansivo. Assim, da Proposição 9, temos

$$\begin{aligned} d(\mathbf{x}^{k+1}, \mathbf{x}^*) &= d(J_{\mu_k}^A(\exp_{\mathbf{x}^k} \mu_k(\mathbf{w}^k + \mathbf{d}^k)), J_{\mu_k}^A(\exp_{\mathbf{x}^*} \mu_k \mathbf{w}^*)) \\ &\leq d(\exp_{\mathbf{x}^k} \mu_k(\mathbf{w}^k + \mathbf{d}^k), \exp_{\mathbf{x}^*} \mu_k \mathbf{w}^*). \end{aligned}$$

Aplicando a desigualdade triangular, segue que

$$\begin{aligned} d(\mathbf{x}^{k+1}, \mathbf{x}^*) &\leq d(\exp_{\mathbf{x}^k} \mu_k(\mathbf{w}^k + \mathbf{d}^k), \mathbf{x}^*) + d(\mathbf{x}^*, \exp_{\mathbf{x}^*} \mu_k \mathbf{w}^*) \\ &\leq d(\mathbf{x}^k, \mathbf{x}^*) + d(\exp_{\mathbf{x}^k} \mu_k(\mathbf{w}^k + \mathbf{d}^k), \mathbf{x}^k) + d(\exp_{\mathbf{x}^*} \mu_k \mathbf{w}^*, \mathbf{x}^*). \end{aligned}$$

Sabemos pela Observação 5, que $d(\exp_{\mathbf{x}^k} \mu_k(\mathbf{w}^k + \mathbf{d}^k), \mathbf{x}^k) = \mu_k \|\mathbf{w}^k + \mathbf{d}^k\|$ e $d(\mathbf{x}^*, \exp_{\mathbf{x}^*} \mu_k \mathbf{w}^*) = \mu_k \|\mathbf{w}^*\|$, então aplicando novamente a desigualdade triangular e em seguida usando a condição (A1) e o fato que a aplicação transporte paralelo é uma isometria, obtemos

$$\begin{aligned} d(\mathbf{x}^{k+1}, \mathbf{x}^*) &= d(\mathbf{x}^k, \mathbf{x}^*) + \mu_k \|\mathbf{w}^k + \mathbf{d}^k\| + \mu_k \|\mathbf{w}^*\| \\ &\leq d(\mathbf{x}^k, \mathbf{x}^*) + \mu_k (\|\mathbf{w}^k - P_{\mathbf{x}^k, \mathbf{x}^*} \mathbf{w}^*\| + \|P_{\mathbf{x}^k, \mathbf{x}^*} \mathbf{w}^*\| + \|\mathbf{d}^k\|) + \mu_k \|\mathbf{w}^*\| \\ &\leq d(\mathbf{x}^k, \mathbf{x}^*) + \mu_k (Ld(\mathbf{x}^k, \mathbf{x}^*) + \|\mathbf{w}^*\| + \|\mathbf{d}^k\|) + \mu_k \|\mathbf{w}^*\| \\ &= (1 + \mu_k L)d(\mathbf{x}^k, \mathbf{x}^*) + \mu_k \gamma_k d(\mathbf{x}^k, \mathbf{x}^{k-1}) + 2\mu_k \|\mathbf{w}^*\|. \end{aligned}$$

Agora aplicando o Lema 1 para as sequências numéricas $\alpha_k = d(\mathbf{x}^k, \mathbf{x}^*)$, $\beta_k = L\mu_k$ e $\Gamma_k = \mu_k \gamma_k d(\mathbf{x}^k, \mathbf{x}^{k-1}) + 2\mu_k \|\mathbf{w}^*\|$ e levando em conta que (A2) e (A3) são válidas, concluímos que $\{d(\mathbf{x}^k, \mathbf{x}^*)\}$ é convergente. Consequentemente $\{\mathbf{x}^k\}$ é limitada. \square

O corolário abaixo apresenta uma importante propriedade, que é um comportamento clássico do método do ponto proximal. A sequência gerado pelo MPPI satisfaz a seguinte propriedade: $\lim_{k \rightarrow +\infty} d(x^{k+1}, x^k) = 0$. O resultado não assume a limitação da sequência ou a monotonicidade forte do campo vetorial B como na Proposição 16.

Corolário 2. *Se as suposições (A1), (A2) e (A3) são válidas, então*

$$\lim_{k \rightarrow +\infty} d(x^{k+1}, x^k) = 0.$$

Demonstração. De (3.11), temos

$$\frac{1}{\mu_k} d^2(x^{k+1}, x^k) \leq \langle u^{k+1}, -\exp_{x^{k+1}}^{-1} x^k \rangle + \langle v^k, -\exp_{x^k}^{-1} x^{k+1} \rangle + \langle d^k, \exp_{x^k}^{-1} x^{k+1} \rangle,$$

$u^{k+1} = B(x^{k+1})$ e $v^k \in A(x^k)$.

Usando o transporte paralelo, temos que $\exp_{x^k}^{-1} x^{k+1} = -P_{x^k, x^{k+1}} \exp_{x^{k+1}}^{-1} x^k$ então,

$$\frac{1}{\mu_k} d^2(x^{k+1}, x^k) \leq \langle P_{x^k, x^{k+1}} u^{k+1} - v^k + d^k, \exp_{x^k}^{-1} x^{k+1} \rangle.$$

Aplicando agora a desigualdade de Cauchy-Schwarz, temos

$$\frac{1}{\mu_k} d^2(x^{k+1}, x^k) \leq \|P_{x^k, x^{k+1}} u^{k+1} - v^k + d^k\| \|\exp_{x^k}^{-1} x^{k+1}\|.$$

Isso implica que

$$d(x^{k+1}, x^k) \leq \mu_k (\|u^{k+1}\| + \|v^k\| + \gamma_k d(x^k, x^{k-1})), \quad (3.19)$$

onde foi usado o fato que $\|\exp_{x^k}^{-1} x^{k+1}\| = d(x^{k+1}, x^k) > 0$.

Segue do Teorema 11, que $\{x^k\}$ é limitada, então pela Proposição 6 e (A3) as sequências $\{u^k\}$, $\{v^k\}$ e $\{\gamma_k d(x^k, x^{k-1})\}$ são limitadas. Portanto, fazendo $k \rightarrow \infty$ em (3.19) e usando (A2), temos que $\lim_{k \rightarrow +\infty} d(x^{k+1}, x^k) = 0$. \square

3.2.3 Condição suficiente para convergência total

Nesta subseção, apresentamos uma condição suficiente para obter a convergência total da sequência gerada pelo MPPI. Sob as suposições (A1), (A2) e (A3) provamos que

$\lim_{k \rightarrow +\infty} d(x^{k+1}, x^k) = 0$ que é um comportamento clássico do método do ponto proximal. Segue de (A2) que $\lim_{k \rightarrow +\infty} \mu_k = 0$. Uma vez que esta sequência auxiliar de parâmetros é escolhida livremente satisfazendo algumas condições adequadas, como feito no capítulo 2 vamos tomá-la de tal forma que ela não vá a zero mais rápido que $d(x^{k+1}, x^k)$. Isso nos leva à seguinte condição:

$$(A4) \text{ Suponha que } \lim_{k \rightarrow \infty} \frac{d(x^{k+1}, x^k)}{\mu_k} = 0.$$

Sabemos pelo Teorema 11 que a sequência gerada pelo MPPI é limitada, ou seja, possui pontos de acumulação. No teorema abaixo, provamos que todo ponto de acumulação da sequência gerada pelo MPPI é um ponto crítico de $A - B$.

Teorema 12. *Se as suposições (A1) – (A4) são válidas, então todo ponto de acumulação de $\{x^k\}$ é ponto crítico de $A - B$.*

Demonstração. Seja \bar{x} um ponto de acumulação arbitrário de $\{x^k\}$, o qual existe pelo Teorema 11. Seja $\{x^{k_j}\}$ uma subsequência de $\{x^k\}$ tal que $x^{k_j} \rightarrow \bar{x}$. Pelo Corolário 2, temos que $\lim_{j \rightarrow \infty} x^{k_j+1} = \bar{x}$. Além disso, por (3.2) temos $w^{k_j} = B(x^{k_j})$, então pela Proposição 6, segue que $\{w^{k_j}\}$ é limitada. Portanto, sem perda de generalidade, podemos (se necessário) tomar uma subsequência $\{w^{k_j}\}$ convergindo para \bar{w} .

Como $w^{k_j} = B(x^{k_j})$ da monotonicidade de B , temos

$$\langle w^{k_j}, \exp_{x^{k_j}}^{-1} z \rangle \leq \langle u, -\exp_z^{-1} x^{k_j} \rangle, \quad \forall u = B(z), z \in M.$$

Fazendo $j \rightarrow \infty$, obtemos

$$\langle \bar{w}, \exp_{\bar{x}}^{-1} z \rangle \leq \langle u, -\exp_z^{-1} \bar{x} \rangle, \quad u = B(z), \forall z \in M.$$

Pela maximalidade de B , segue que $\bar{w} = B(\bar{x})$.

Por outro lado, segue de (3.3) que

$$\frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} y^{k_j} \in A(x^{k_j+1}),$$

onde $y^{k_j} = \exp_{x^{k_j}} \mu_{k_j} (w^{k_j} + d^{k_j})$.

Afirmamos que $\lim_{j \rightarrow \infty} \frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} = \bar{\mathbf{w}}$, e assim, $\bar{\mathbf{w}} \in \mathcal{A}(\bar{\mathbf{x}})$.

Com efeito,

$$\begin{aligned} & \left\| \frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} - \mathbf{P}_{x^{k_j+1}, \bar{\mathbf{x}}} \bar{\mathbf{w}} \right\| \leq \\ & \left\| \frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} - \mathbf{P}_{x^{k_j+1}, x^{k_j}}(\mathbf{w}^{k_j} + \mathbf{d}^{k_j}) \right\| + \left\| \mathbf{P}_{x^{k_j+1}, x^{k_j}}(\mathbf{w}^{k_j} + \mathbf{d}^{k_j}) - \mathbf{P}_{x^{k_j+1}, \bar{\mathbf{x}}} \bar{\mathbf{w}} \right\|. \end{aligned} \quad (3.20)$$

Vamos provar que o lado direito da desigualdade acima se anula quando $j \rightarrow \infty$. Primeiro, note que pelo fato que a aplicação transporte paralelo é uma isometria e usando que $\mathbf{w}^{k_j} + \mathbf{d}^{k_j} = \frac{1}{\mu_{k_j}} \exp_{x^{k_j}}^{-1} \mathbf{y}^{k_j}$, temos para o primeiro termo que

$$\begin{aligned} & \left\| \frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} - \mathbf{P}_{x^{k_j+1}, x^{k_j}}(\mathbf{w}^{k_j} + \mathbf{d}^{k_j}) \right\|^2 = \\ & = \frac{1}{\mu_{k_j}^2} \left\| \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} \right\|^2 + \frac{1}{\mu_{k_j}^2} \left\| \exp_{x^{k_j}}^{-1} \mathbf{y}^{k_j} \right\|^2 - 2 \frac{1}{\mu_{k_j}} \left\langle \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j}, \mathbf{P}_{x^{k_j+1}, x^{k_j}}(\mathbf{w}^{k_j} + \mathbf{d}^{k_j}) \right\rangle \\ & = \frac{1}{\mu_{k_j}^2} \left(\mathbf{d}^2(x^{k_j+1}, \mathbf{y}^{k_j}) + \mathbf{d}^2(\mathbf{y}^{k_j}, x^{k_j}) \right) - 2 \frac{1}{\mu_{k_j}} \left\langle \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j}, \mathbf{P}_{x^{k_j+1}, x^{k_j}}(\mathbf{w}^{k_j} + \mathbf{d}^{k_j}) \right\rangle. \end{aligned}$$

Novamente pela isometria do transporte paralelo e as seguintes igualdades

$\mathbf{P}_{\mathbf{y}^{k_j}, x^{k_j+1}}(\mathbf{P}_{x^{k_j+1}, x^{k_j}}) = \mathbf{P}_{\mathbf{y}^{k_j}, x^{k_j}}$ e $\mathbf{P}_{\mathbf{y}^{k_j}, x^{k_j+1}}(\exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j}) = -\exp_{\mathbf{y}^{k_j}}^{-1} x^{k_j+1}$, segue que

$$\begin{aligned} & \left\| \frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} - \mathbf{P}_{x^{k_j+1}, x^{k_j}}(\mathbf{w}^{k_j} + \mathbf{d}^{k_j}) \right\|^2 = \\ & = \frac{1}{\mu_{k_j}^2} \left(\mathbf{d}^2(x^{k_j+1}, \mathbf{y}^{k_j}) + \mathbf{d}^2(\mathbf{y}^{k_j}, x^{k_j}) \right) \\ & \quad - \frac{2}{\mu_{k_j}} \left\langle \mathbf{P}_{\mathbf{y}^{k_j}, x^{k_j+1}}(\exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j}), \mathbf{P}_{\mathbf{y}^{k_j}, x^{k_j+1}}(\mathbf{P}_{x^{k_j+1}, x^{k_j}}(\mathbf{w}^{k_j} + \mathbf{d}^{k_j})) \right\rangle \\ & = \frac{1}{\mu_{k_j}^2} \left(\mathbf{d}^2(x^{k_j+1}, \mathbf{y}^{k_j}) + \mathbf{d}^2(\mathbf{y}^{k_j}, x^{k_j}) \right) - \frac{2}{\mu_{k_j}} \left\langle -\exp_{\mathbf{y}^{k_j}}^{-1} x^{k_j+1}, \mathbf{P}_{\mathbf{y}^{k_j}, x^{k_j}} \left(\frac{1}{\mu_{k_j}} \exp_{x^{k_j}}^{-1} \mathbf{y}^{k_j} \right) \right\rangle. \end{aligned}$$

Como o transporte paralelo uma isometria linear, obtemos que

$$\mathbf{P}_{\mathbf{y}^{k_j}, x^{k_j}} \left(\frac{1}{\mu_{k_j}} \exp_{x^{k_j}}^{-1} \mathbf{y}^{k_j} \right) = \frac{1}{\mu_{k_j}} \mathbf{P}_{\mathbf{y}^{k_j}, x^{k_j}} \left(\exp_{x^{k_j}}^{-1} \mathbf{y}^{k_j} \right) = -\frac{1}{\mu_{k_j}} \exp_{\mathbf{y}^{k_j}}^{-1} x^{k_j},$$

implicando que

$$\begin{aligned}
 & \left\| \frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} - \mathbf{P}_{x^{k_j+1}, x^{k_j}}(\mathbf{w}^{k_j} + \mathbf{d}^{k_j}) \right\|^2 = \\
 &= \frac{1}{\mu_{k_j}^2} \left(\mathbf{d}^2(x^{k_j+1}, \mathbf{y}^{k_j}) + \mathbf{d}^2(\mathbf{y}^{k_j}, x^{k_j}) \right) - 2 \frac{1}{\mu_{k_j}} \left\langle -\exp_{y^{k_j}}^{-1} x^{k_j+1}, -\frac{1}{\mu_{k_j}} \exp_{y^{k_j}}^{-1} x^{k_j} \right\rangle \\
 &= \frac{1}{\mu_{k_j}^2} \left(\mathbf{d}^2(x^{k_j+1}, \mathbf{y}^{k_j}) + \mathbf{d}^2(\mathbf{y}^{k_j}, x^{k_j}) \right) - 2 \frac{1}{\mu_{k_j}^2} \left\langle \exp_{y^{k_j}}^{-1} x^{k_j+1}, \exp_{y^{k_j}}^{-1} x^{k_j} \right\rangle. \quad (3.21)
 \end{aligned}$$

Pelo triângulo geodésico $\Delta(x^{k_j+1}, \mathbf{y}^{k_j}, x^{k_j})$, temos

$$\mathbf{d}^2(x^{k_j+1}, \mathbf{y}^{k_j}) + \mathbf{d}^2(\mathbf{y}^{k_j}, x^{k_j}) - 2 \langle \exp_{y^{k_j}}^{-1} x^{k_j+1}, \exp_{y^{k_j}}^{-1} x^{k_j} \rangle \leq \mathbf{d}^2(x^{k_j+1}, x^{k_j}).$$

Usando a última desigualdade em (3.21), temos

$$\left\| \frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} - \mathbf{P}_{x^{k_j+1}, x^{k_j}}(\mathbf{w}^{k_j} + \mathbf{d}^{k_j}) \right\| \leq \frac{1}{\mu_{k_j}} \mathbf{d}(x^{k_j+1}, x^{k_j}).$$

Fazendo $j \rightarrow \infty$ na desigualdade acima e usando (A4), obtemos

$$\lim_{j \rightarrow \infty} \left\| \frac{1}{\mu_{k_j}} \exp_{x^{k_j+1}}^{-1} \mathbf{y}^{k_j} - \mathbf{P}_{x^{k_j+1}, x^{k_j}}(\mathbf{w}^{k_j} + \mathbf{d}^{k_j}) \right\| = 0. \quad (3.22)$$

Agora, usando o fato que $\mathbf{P}_{x^{k_j+1}, \bar{x}} = \mathbf{P}_{x^{k_j+1}, x^{k_j}}(\mathbf{P}_{x^{k_j}, \bar{x}})$, a isometria do transporte paralelo, a desigualdade triangular, (A1) e $\|\mathbf{d}^{k_j}\| = \|\gamma_{k_j} \exp_{x^{k_j}}^{-1} x^{k_j-1}\| = \gamma_{k_j} \mathbf{d}(x^{k_j}, x^{k_j-1})$, temos que

$$\begin{aligned}
 \|\mathbf{P}_{x^{k_j+1}, x^{k_j}}(\mathbf{w}^{k_j} + \mathbf{d}^{k_j}) - \mathbf{P}_{x^{k_j+1}, \bar{x}} \bar{\mathbf{w}}\| &= \|\mathbf{P}_{x^{k_j+1}, x^{k_j}}(\mathbf{w}^{k_j} + \mathbf{d}^{k_j}) - \mathbf{P}_{x^{k_j+1}, x^{k_j}}(\mathbf{P}_{x^{k_j}, \bar{x}} \bar{\mathbf{w}})\| \\
 &= \|\mathbf{w}^{k_j} + \mathbf{d}^{k_j} - \mathbf{P}_{x^{k_j}, \bar{x}} \bar{\mathbf{w}}\| \\
 &\leq \|\mathbf{w}^{k_j} - \mathbf{P}_{x^{k_j}, \bar{x}} \bar{\mathbf{w}}\| + \|\mathbf{d}^{k_j}\| \\
 &\leq L \mathbf{d}(x^{k_j}, \bar{x}) + \gamma_{k_j} \mathbf{d}(x^{k_j}, x^{k_j-1}).
 \end{aligned}$$

Pelo Corolário 2, temos que $\lim_{j \rightarrow \infty} \gamma_{k_j} \mathbf{d}(x^{k_j}, x^{k_j-1}) = 0$. Assim, fazendo $j \rightarrow \infty$ na

desigualdade acima e usando o fato que $\mathbf{x}^{k_j} \rightarrow \bar{\mathbf{x}}$, obtemos

$$\lim_{j \rightarrow \infty} \|\mathbf{P}_{\mathbf{x}^{k_j+1}, \mathbf{x}^{k_j}}(\mathbf{w}^{k_j} + \mathbf{d}^{k_j}) - \mathbf{P}_{\mathbf{x}^{k_j+1}, \bar{\mathbf{x}}}\bar{\mathbf{w}}\| = 0. \quad (3.23)$$

Fazendo $j \rightarrow \infty$ em (3.20) e, usando (3.22) e (3.23), temos

$$\lim_{j \rightarrow \infty} \frac{1}{\mu_{k_j}} \exp_{\mathbf{x}^{k_j+1}}^{-1} \mathbf{y}^{k_j} = \bar{\mathbf{w}}.$$

Usando agora a monotonicidade de \mathbf{A} , temos que

$$\left\langle \frac{1}{\mu_{k_j}} \exp_{\mathbf{x}^{k_j+1}}^{-1} \mathbf{y}^{k_j}, \exp_{\mathbf{x}^{k_j+1}}^{-1} \mathbf{z} \right\rangle \leq \langle \mathbf{u}, -\exp_{\mathbf{z}}^{-1} \mathbf{x}^{k_j+1} \rangle, \quad \forall \mathbf{u} \in \mathbf{A}(\mathbf{z}), \mathbf{z} \in \mathbf{M}.$$

Fazendo $j \rightarrow \infty$, obtemos

$$\langle \bar{\mathbf{w}}, \exp_{\bar{\mathbf{x}}}^{-1} \mathbf{z} \rangle \leq \langle \mathbf{u}, -\exp_{\mathbf{z}}^{-1} \bar{\mathbf{x}} \rangle, \quad \forall \mathbf{u} \in \mathbf{A}(\mathbf{z}), \mathbf{z} \in \mathbf{M}.$$

Pela maximalidade de \mathbf{A} , segue que $\bar{\mathbf{w}} \in \mathbf{A}(\bar{\mathbf{x}})$, e portanto, $\bar{\mathbf{w}} \in \mathbf{A}(\bar{\mathbf{x}}) \cap \mathbf{B}(\bar{\mathbf{x}})$ implicando que $\bar{\mathbf{x}}$ é um ponto crítico de $\mathbf{A} - \mathbf{B}$, isto é, $\bar{\mathbf{x}} \in \mathbf{S}$. \square

Mostramos no teorema abaixo que a sequência gerada pelo MPPI converge para um ponto crítico de $\mathbf{A} - \mathbf{B}$.

Teorema 13. *Se as suposições (A1) – (A4) são válidas, então $\{\mathbf{x}^k\}$ converge para um ponto crítico de $\mathbf{A} - \mathbf{B}$.*

Demonstração. Com efeito, segue do Teorema 11 que a sequência $\{\mathbf{x}^k\}$ é limitada, então existe $\{\mathbf{x}^{k_j}\}$ subsequência de $\{\mathbf{x}^k\}$ tal que $\lim_{j \rightarrow \infty} \mathbf{x}^{k_j} = \mathbf{x}^*$. Assim,

$$\lim_{j \rightarrow \infty} \mathbf{d}(\mathbf{x}^{k_j}, \mathbf{x}^*) = 0.$$

Como \mathbf{x}^* é um ponto de acumulação de $\{\mathbf{x}^k\}$, segue do Teorema 12 que $\mathbf{x}^* \in \mathbf{S}$. Então, pelo Teorema 11, a sequência $\{\mathbf{d}(\mathbf{x}^k, \mathbf{x}^*)\}$ é convergente implicando que

$$\lim_{k \rightarrow \infty} \mathbf{d}(\mathbf{x}^k, \mathbf{x}^*) = 0.$$

Portanto, $\{\mathbf{x}^k\}$ converge para $\mathbf{x}^* \in \mathbf{S}$, ou seja, $\{\mathbf{x}^k\}$ converge para um ponto crítico de $\mathbf{A} - \mathbf{B}$. \square

3.3 Análise de convergência: caso monótono

Ressaltamos que o MPPI é novo, no cenário de Hadamard, mesmo para resolver campos vetoriais monótonos. Portanto, consideramos uma versão monótona do MPPI fazendo $B(x) = 0$. Então, o MPPI se torna o seguinte:

MPPI: versão monótona - mMPPI

Passo 1: Dados um ponto inicial $x^0 \in M$, $\gamma_k \in [0, \frac{\rho}{2})$ e considere $a, b > 0$. Seja $\{\mu_k\}$ uma sequência de números positivos tal que $0 < a \leq \mu_k \leq b$. Defina $x^{-1} = x^0$.

Passo 2: Dado $x^k \in M$, defina $d^k = \gamma_k \exp_{x^k}^{-1} x^{k-1}$ e

$$y^k := \exp_{x^k} \mu_k d^k. \quad (3.24)$$

Passo 3: Calcule $x^{k+1} \in M$ tal que

$$0 \in A(x^{k+1}) - \frac{1}{\mu_k} \exp_{x^{k+1}}^{-1} y^k. \quad (3.25)$$

Se $x^{k+1} = x^k$ e $d^k = 0$, pare. Caso contrário, defina $k := k + 1$ e retorne ao Passo 2.

O objetivo desta seção é estudar as propriedades de convergência do mMPPI. Como o problema agora é monótono é natural que obtenhamos os mesmos resultados que na versão possivelmente não monótona sob suposições menos restritivas.

Para provar a convergência do mMPPI vamos considerar $A^{-1}(0) \subset M$ o conjunto de singularidades de A e assumir que $A^{-1}(0) \neq \emptyset$, ou seja,

$$A^{-1}(0) = \{x \in M; 0 \in A(x)\} \neq \emptyset.$$

Nesta seção, também consideramos a seguinte condição:

$$\sum_{k=0}^{\infty} \gamma_k d(x^k, x^{k-1}) < \infty. \quad (3.26)$$

Observação 23. *Em espaços de Hilbert, (ALIMOHAMMADY; RAMAZANNEJAD, 2016; KHATIBZADEH; RANJBAR, 2015) consideram (3.26), isto é, $\sum_{k=0}^{\infty} \gamma_k \|x^k - x^{k-1}\| < \infty$. Uma suposição semelhante foi considerada, por exemplo, em (ALVAREZ; ATTOUCH, 2001, Teorema 2.1), onde assume-se que $\sum_{k=0}^{\infty} \gamma_k \|x^k - x^{k-1}\|^2 < \infty$.*

Os próximos resultados garantem que a sequência $\{x^k\}$ gerada pelo mMPPI converge

para uma singularidade de A .

Teorema 14. *A sequência $\{d(x^k, x^*)\}$ é convergente para cada $x^* \in A^{-1}(0)$ e, em particular, $\{x^k\}$ é limitada.*

Demonstração. Por (3.25), temos que $\exp_{x^{k+1}}^{-1} y^k \in \mu_k A(x^{k+1})$ e conseqüentemente

$$y^k \in \exp_{x^{k+1}} \mu_k A(x^{k+1}).$$

Isso significa que $x^{k+1} = J_{\mu_k}^A(y^k)$ e por (3.24), temos $y^k = \exp_{x^k} \mu_k d^k$ implicando que

$$x^{k+1} = J_{\mu_k}^A(\exp_{x^k}(\mu_k d^k)).$$

Seja $x^* \in A^{-1}(0)$ fixado, então $0 \in A(x^*)$ tal que

$$x^* = J_{\mu_k}^A(\exp_{x^*} \mu_k 0) = J_{\mu_k}^A(x^*).$$

Como A é monótono maximal, pela Proposição 10, seu resolvente é firmemente não expansivo. Assim, pela Proposition 9, temos

$$\begin{aligned} d(x^{k+1}, x^*) &= d(J_{\mu_k}^A(\exp_{x^k} \mu_k d^k), J_{\mu_k}^A(x^*)) \\ &\leq d(\exp_{x^k} \mu_k d^k, x^*) \\ &\leq d(x^k, x^*) + d(\exp_{x^k} \mu_k d^k, x^k) \\ &= d(x^k, x^*) + \mu_k \|d^k\| \\ &= d(x^k, x^*) + \mu_k \gamma_k d(x^k, x^{k-1}), \end{aligned}$$

onde a desigualdade triangular foi aplicada e o fato que M é uma variedade de Hadamard, ou seja, $d(\exp_{x^k} \mu_k d^k, x^k) = \|\exp_{x^k}^{-1}(\exp_{x^k} \mu_k d^k)\| = \mu_k \|d^k\|$.

Portanto, aplicando o Lema 1 com $\alpha_k = d(x^k, x^*)$, $\beta_k = 0$, $\Gamma_k = \mu_k \gamma_k d(x^k, x^{k-1})$ e levando em conta que $\sum_{k=0}^{\infty} \gamma_k d(x^k, x^{k-1}) < \infty$ e $\{\mu_k\}$ é limitada, segue que $\{d(x^k, x^*)\}$ é convergente. Conseqüentemente $\{x^k\}$ é limitada. \square

Corolário 3. *A afirmação a seguir é válida*

$$\lim_{k \rightarrow +\infty} d(x^{k+1}, x^k) = 0.$$

Demonstração. Por (3.25), temos que $\frac{1}{\mu_k} \exp_{x^{k+1}}^{-1} y^k \in A(x^{k+1})$ e seja $x^* \in A^{-1}(0)$, consequentemente pela monotonicidade de A temos que

$$\left\langle \frac{1}{\mu_k} \exp_{x^{k+1}}^{-1} y^k, \exp_{x^{k+1}}^{-1} x^* \right\rangle \leq \langle 0, -\exp_{x^*}^{-1} x^{k+1} \rangle = 0. \quad (3.27)$$

Seja $\triangle(y^k, x^{k+1}, x^*)$ o triângulo geodésico com $\theta = \angle(\exp_{x^{k+1}}^{-1} y^k, \exp_{x^{k+1}}^{-1} x^*)$, então pelo Teorema 3

$$d^2(y^k, x^{k+1}) + d^2(x^{k+1}, x^*) - 2\langle \exp_{x^{k+1}}^{-1} y^k, \exp_{x^{k+1}}^{-1} x^* \rangle \leq d^2(y^k, x^*). \quad (3.28)$$

Sendo $\mu_k > 0$, segue de (3.27) e (3.28) que

$$d^2(y^k, x^{k+1}) + d^2(x^{k+1}, x^*) \leq d^2(y^k, x^*). \quad (3.29)$$

Pela desigualdade triangular, temos que

$$d^2(y^k, x^*) \leq [d(y^k, x^k) + d(x^k, x^*)]^2 = d^2(y^k, x^k) + 2d(y^k, x^k)d(x^k, x^*) + d^2(x^k, x^*), \quad (3.30)$$

então por (3.29) e (3.30), segue que

$$d^2(y^k, x^{k+1}) \leq d^2(y^k, x^k) + 2d(y^k, x^k)d(x^k, x^*) + [d^2(x^k, x^*) - d^2(x^{k+1}, x^*)]. \quad (3.31)$$

Sendo $y^k = \exp_{x^k} \mu_k d^k$, temos $d(y^k, x^k) = \mu_k \gamma_k d(x^k, x^{k-1})$ então $\lim_{k \rightarrow \infty} d(y^k, x^k) = 0$. Assim, fazendo $k \rightarrow \infty$ em (3.31) e usando o fato que $\{d(x^k, x^*)\}$ é convergente, obtemos $\lim_{k \rightarrow \infty} d(y^k, x^{k+1}) = 0$. Então, usando desigualdade triangular, temos que

$$\lim_{k \rightarrow \infty} d(x^{k+1}, x^k) \leq \lim_{k \rightarrow \infty} [d(x^{k+1}, y^k) + d(y^k, x^k)] = 0.$$

Portanto, $\lim_{k \rightarrow \infty} d(x^{k+1}, x^k) = 0$. □

Teorema 15. *A seqüência $\{x^k\}$ gerada pelo mMPPI converge para uma singularidade de A .*

Demonstração. Em virtude do Teorema 14 basta mostrar que todo ponto de acumulação de $\{x^k\}$ pertence a $A^{-1}(0)$. De fato, se x^* é um ponto de acumulação de $\{x^k\}$ que pertence a $A^{-1}(0)$, então existe uma subsequencia $\{x^{k_i}\}$ convergindo para x^* , logo $\{d(x^{k_i}, x^*)\}$

converge a zero. Como pelo Teorema 14, temos que $\{\mathbf{d}(\mathbf{x}^k, \mathbf{x}^*)\}$ é convergente, então $\{\mathbf{d}(\mathbf{x}^k, \mathbf{x}^*)\}$ converge para zero, e consequentemente, $\{\mathbf{x}^k\}$ converge para $\mathbf{x}^* \in \mathbf{A}^{-1}(0)$.

Agora, provaremos que um ponto de acumulação arbitrário $\bar{\mathbf{x}}$ de $\{\mathbf{x}^k\}$ pertence a $\mathbf{A}^{-1}(0)$. Seja $\{\mathbf{x}^{k_j}\}$ uma subsequência de $\{\mathbf{x}^k\}$ tal que $\lim_{j \rightarrow \infty} \mathbf{x}^{k_j} = \bar{\mathbf{x}}$. Pelo Corolário 3, temos que $\lim_{j \rightarrow \infty} \mathbf{x}^{k_j+1} = \bar{\mathbf{x}}$.

Como $\mathbf{y}^k = \exp_{\mathbf{x}^k} \mu_k \mathbf{d}^k$, temos que $\mathbf{d}(\mathbf{y}^k, \mathbf{x}^k) = \mu_k \gamma_k \mathbf{d}(\mathbf{x}^k, \mathbf{x}^{k-1})$, então $\lim_{k \rightarrow \infty} \mathbf{d}(\mathbf{y}^k, \mathbf{x}^k) = 0$. Assim, $\lim_{j \rightarrow +\infty} \mathbf{d}(\mathbf{y}^{k_j}, \mathbf{x}^{k_j}) = 0$, então $\lim_{j \rightarrow \infty} \mathbf{y}^{k_j} = \bar{\mathbf{x}}$. Pelo item (i) da Proposição 1, segue que $\lim_{j \rightarrow +\infty} \exp_{\mathbf{x}^{k_j+1}}^{-1} \mathbf{y}^{k_j} = \exp_{\bar{\mathbf{x}}}^{-1} \bar{\mathbf{x}} = 0$. Sendo $\{\mu_{k_j}\}$ limitada, temos que

$$\lim_{j \rightarrow +\infty} \frac{1}{\mu_{k_j}} \exp_{\mathbf{x}^{k_j+1}}^{-1} \mathbf{y}^{k_j} = 0.$$

Usando agora a monotonicidade de \mathbf{A} para $\frac{1}{\mu_{k_j}} \exp_{\mathbf{x}^{k_j+1}}^{-1} \mathbf{y}^{k_j} \in \mathbf{A}(\mathbf{x}^{k_j+1})$, temos

$$\left\langle \frac{1}{\mu_{k_j}} \exp_{\mathbf{x}^{k_j+1}}^{-1} \mathbf{y}^{k_j}, \exp_{\mathbf{x}^{k_j+1}}^{-1} \mathbf{z} \right\rangle \leq \langle \mathbf{u}, -\exp_{\mathbf{z}}^{-1} \mathbf{x}^{k_j+1} \rangle, \quad \forall \mathbf{u} \in \mathbf{A}(\mathbf{z}), \mathbf{z} \in \mathbf{M}.$$

Fazendo $j \rightarrow +\infty$ na última desigualdade e usando a Proposição 1, temos que

$$\langle 0, \exp_{\bar{\mathbf{x}}}^{-1} \mathbf{z} \rangle \leq \langle \mathbf{u}, -\exp_{\mathbf{z}}^{-1} \bar{\mathbf{x}} \rangle, \quad \forall \mathbf{u} \in \mathbf{A}(\mathbf{z}), \mathbf{z} \in \mathbf{M}.$$

Portanto, a maximalidade de \mathbf{A} implica que $0 \in \mathbf{A}(\bar{\mathbf{x}})$, isto é, $\bar{\mathbf{x}} \in \mathbf{A}^{-1}(0)$. □

Capítulo 4

Experimentos numéricos

Neste capítulo, realizamos alguns experimentos numéricos para ilustrar nossos resultados teóricos. Usamos os métodos do ponto proximal para resolver diferentes problemas.

Os experimentos numéricos são codificados em MATLAB R2020b em uma máquina com Intel(R) Core(TM) i7, CPU de 2,3 GHz e 8 GB de memória. Consideramos exemplos em uma variedade de Hadamard genuína (com curvatura diferente de zero) e no cenário Euclidiano. No contexto de Hadamard, usamos o pacote para otimização em variedades "MANOPT"; veja (BOUMAL et al., 2014). O critério de parada usado foi $d(x^{k+1}, x^k) < \epsilon$ ou o número de iterações atingir "maxiter". Este critério também é usado no contexto Euclidiano onde a distância Riemanniana é a norma Euclidiana. Para cada exemplo abaixo, executamos o algoritmo 100 vezes usando diferentes pontos de partida aleatórios e consideramos diferentes dimensões. Em todos os exemplos, comparamos o desempenho do método com diferentes escolhas do parâmetro μ_k .

Antes de apresentar os exemplos, vamos definir a variedade de Hadamard que consideraremos na sequência. Referimos a (ABSIL et al., 2009) para mais detalhes.

Sejam \mathbb{P}^n o conjunto das matrizes simétricas, \mathbb{P}_+^n o cone das matrizes simétricas semi-definidas positivas e \mathbb{P}_{++}^n o cone das matrizes simétricas definidas positivas $n \times n$. Seja $M := (\mathbb{P}_{++}^n, \langle \cdot, \cdot \rangle)$ a variedade Riemanniana dotada da métrica Riemanniana dada por

$$\langle \mathbf{u}, \mathbf{v} \rangle_X := \text{tr}(\mathbf{V}\mathbf{X}^{-1}\mathbf{u}\mathbf{X}^{-1}), \quad \mathbf{X} \in M, \quad \mathbf{u}, \mathbf{v} \in T_X M,$$

onde $\text{tr}(\mathbf{X})$ denota o traço de $\mathbf{X} \in \mathbb{P}^n$ e $T_X M \approx \mathbb{P}^n$, com a correspondente norma denotada por $\| \cdot \|$; veja (ROTHAUS, 1960). Neste caso, para quaisquer $\mathbf{X}, \mathbf{Y} \in M$ a única geodésica

ligando esses dois pontos é dada por:

$$\gamma(t) = X^{1/2} (X^{-1/2} Y X^{-1/2})^t X^{1/2}, \gamma'(0) = X^{1/2} \ln(X^{-1/2} Y X^{-1/2}) X^{1/2},$$

$t \in [0, 1]$, veja, por exemplo, (NESTEROV; TODD, 2002) e (BENTO et al., 2015). Mais precisamente, M é uma variedade de Hadamard, veja, por exemplo, (LANG, 2012, Teorema 1.2., página 325). Pode-se calcular a curvatura de M e verificar que ela possui curvatura não constante; veja (LENGLET et al., 2006, página 428).

A aplicação exponencial $\exp_X : T_X M \rightarrow M$ e sua inversa $\exp_X^{-1} : M \rightarrow T_X M$ são dadas da seguinte forma

$$\exp_X V = X^{1/2} e^{X^{-1/2} V X^{-1/2}} X^{1/2}, \quad \exp_X^{-1} Y = X^{1/2} \ln(X^{-1/2} Y X^{-1/2}) X^{1/2},$$

veja (BHATIA, 2009, Capítulo 6) e (BOUMAL, 2022, Capítulo 11).

Neste contexto, para gerar as matrizes aleatórias usamos o comando "manifold.rand". Os subproblemas são resolvidos usando a rotina "steepestdescent" com o loop interno parando se a norma do gradiente for menor que $\epsilon = 10^{-5}$ ou o número de iterações atingir "maxiter = 1000".

Problema 1. *Consideramos o problema de minimização convexa*

$$\min_{X \in M} f(X),$$

onde $f(X) = \frac{1}{2} \ln(\det(X))^2$ é convexa conforme item (i) no Exemplo 4, em M , porém verifica-se que f não é convexa no sentido Euclidiano. Claramente, este problema pode ser visto como $0 \in A(X) - B(X)$ para $A(X) = \text{grad} f(X)$ e $B(X) = 0$, onde A é um campo vetorial monótono maximal pelos Exemplos 4 e 11. A solução da equação $0 = \text{grad} f(X)$ é uma matriz X^* tal que $\det(X^*) = 1$ e, portanto, $f(X^*) = 0$. Segue pelo Exemplo 4 que $\text{grad} f(X) = \ln(\det(X))X$ e $\text{hess} f(X)V = \text{tr}(X^{-1}V)X$.

Problema 2. *Consideramos o problema DC não convexo*

$$\min_{X \in M} f(X) = g(X) - h(X),$$

onde $f(X) = \ln(\det(X))^4 - 2 \ln(\det(X))^2 + 1$ com $g(X) = \ln(\det(X))^4 + 1$ e

$h(X) = 2 \ln(\det(X))^2$ funções convexas conforme item (i) e (ii) no Exemplo 4. Vale ressaltar que g e h não são funções convexas no sentido Euclidiano. Apesar da convexidade de g e h em M , temos que f é uma função não-convexa em ambas as configurações Euclidiana e Riemanniana. Claramente, este problema pode ser visto como $0 \in A(X) - B(X)$ para $A(X) = \text{grad } g(X)$ e $B(X) = \text{grad } h(X)$, onde A e B são campos vetoriais monótonos maximais pelos Exemplos 4 e 11. A solução da equação $0 = \text{grad } f(X)$ é uma matriz X^* tal que $\det(X^*) = 1$, $\det(X^*) = e$ ou $\det(X^*) = e^{-1}$ com $\min_{X \in M} f(X) = 0$.

A seguir, consideramos alguns problemas onde o campo vetorial é definido no espaço Euclidiano incluindo casos em que o campo vetorial não é o subdiferencial (ou o gradiente) de uma função convexa.

Problema 3. *Sejam os operadores $A, B : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dados por $A(x) = \left(\frac{x_1}{2} - x_2, x_1 + \frac{x_2}{2} \right)$ e $B(x) = (x_2, -x_1)$ com $x = (x_1, x_2)$. Temos que A é fortemente monótono e maximal e B é monótono maximal mas nenhum deles é o subdiferencial de uma função convexa porque não são operadores ciclicamente monótonos; veja (ROCKAFELLAR, 1966, Teorema 3). Considere o problema de inclusão variacional*

$$0 \in C(x) = A(x) - B(x)$$

que é equivalente a mostrar que $A(x) \cap B(x)$ é não vazio, para algum $x \in \mathbb{R}^2$. Podemos facilmente mostrar que tomando $x^* = (0, 0)$, temos que $0 \in A(x^*) \cap B(x^*)$. Assim, $x^* = (0, 0)$ é uma solução da inclusão variacional acima.

Problema 4. *Consideramos o problema DC não convexo*

$$\min_{(x \in \mathbb{R}^2)} f(x) = g(x) - h(x),$$

onde $x = (x_1, x_2)$, $f(x) = \frac{1}{2} \|x\|^2 + \|x\| - \max\{-x_1, 0\} - \max\{-x_2, 0\}$ com $g(x) = \|x\|^2 + \|x\|$ e $h(x) = \frac{1}{2} \|x\|^2 + \max\{-x_1, 0\} + \max\{-x_2, 0\}$ funções fortemente convexas. Claramente, este problema pode ser visto como $0 \in A(X) - B(X)$ para $A(x) = \partial g(x)$ e $B(x) = \partial h(x)$, onde A e B são operadores monótonos maximais (ponto conjunto); veja Exemplo 11.

Comparamos o desempenho do MPP e MPPI para diferentes parâmetros μ_k e γ_k em várias instâncias. Nos experimentos iremos denotar por MPP o caso do MPPI com $\gamma_k = 0$,

para todo $k \in \mathbb{N}$ conforme mencionado anteriormente. Os resultados são apresentados na Tabela 4.1 para o Problema 1, Tabela 4.2 para o Problema 2, Tabela 4.4 para o Problema 3 e Tabela 4.3 para o Problema 4. Nestas tabelas, a coluna n denota a dimensão, μ_k e γ_k mostram as escolhas destes parâmetros em MPP e MPPI, $\text{min.iter.}(k)$ (resp. min. time(s)) e max. iter. (resp. max. time(s)) são o número mínimo e o máximo de iterações (resp. tempo de CPU em segundos) até que o método satisfaça a regra de parada, respectivamente, e med. iter. (resp. med. time(s)) representa a média de iterações (resp. tempo de CPU) em 100 execuções.

Nas tabelas, apresentamos os resultados do MPP e MPPI para diferentes valores de μ_k e γ_k . Mais precisamente, consideramos para os Problemas 1, 2 e 3 valores de μ_k constantes satisfazendo as condições 2.4 e 2.21 para o MPP e as condições 3.1 e 3.18 para o MPPI, além de valores de μ_k variável satisfazendo a hipótese (A2) para ambos os métodos. Como o Problema 4 não satisfaz a condição (A1), para esse problema consideramos apenas valores de μ_k constante. Em todos os problemas consideramos diferentes valores de γ_k constantes para ambos os métodos.

As Figuras 4.1, 4.2, 4.3, 4.4 para o Problema 1, Figuras 4.5, 4.6, 4.7, 4.8 para o Problema 2, Figura 4.9 para o Problema 3 e Figura 4.10 para o Problema 4 ilustram o comportamento dos métodos MPP e MPPI em relação aos valores $\frac{d(x^{k+1}, x^k)}{\mu_k}$, $\|d^k\|$ e $|f(x^k)|$ (ou $\|C(x^k)\|$). Essas figuras comprovam que os exemplos considerados satisfazem algumas hipótese e resultados apresentados anteriormente.

Nas figuras apresentamos os resultados de uma rodada do método MPP e três variações do método MPPI a saber MPPI 1, MPPI 2 e MPPI-v. As variações do MPPI depende dos valores de μ_k escolhido. Os valores de μ_k e γ_k usados no MPP, MPPI 1, MPPI 2 e MPPI-v são respectivamente $\mu_k = 0.1$ e $\gamma_k = 0$, $\mu_k = 5$ e $\gamma_k = 0.01$, $\mu_k = 1$ e $\gamma_k = 0.1$, $\mu_k = \frac{10}{(k+1)^2}$ e $\gamma_k = 0.1$ para o Problema 1. Os valores de μ_k e γ_k usados no MPP, MPPI 1, MPPI 2 e MPPI-v são respectivamente $\mu_k = 1$ e $\gamma_k = 0$, $\mu_k = 0.1$ e $\gamma_k = 0.01$, $\mu_k = 0.5$ e $\gamma_k = 0.1$, $\mu_k = \frac{2}{(k+1)^2}$ e $\gamma_k = 0.1$ para o Problema 2. Os valores de μ_k e γ_k usados no MPP, MPPI 1, MPPI 2 e MPPI-v são respectivamente $\mu_k = 0.1$ e $\gamma_k = 0$, $\mu_k = 0.5$ e $\gamma_k = 0.1$, $\mu_k = 1$ e $\gamma_k = 0.5$, $\mu_k = \frac{10^3}{(k+1)^2}$ e $\gamma_k = 0.5$ para o Problema 3. Para o Problema 4, consideramos apenas MPP, MPPI 1 e MPPI 2 com valores de $\mu_k = 1$ e $\gamma_k = 0$, $\mu_k = 5$ e $\gamma_k = 0.1$, $\mu_k = 1$ e $\gamma_k = 0.01$, respectivamente.

Tabela 4.1: Algoritmos MPP (FERREIRA; OLIVEIRA, 2002) e mMPPi para Problema 1.

| n | μ_k | γ_k | min. iter.(k) | max. iter.(k) | med. iter.(k) | min. time | max. time | med. time | grad f(X ^k) | f(X ^k) |
|-----|----------------------|------------|---------------|---------------|---------------|-----------|-----------|-----------|-------------------------|--------------------|
| 5 | 0.1 | 0 | 25 | 28 | 26.67 | 0.31162 | 1.3126 | 0.36525 | 8.4047e-05 | 7.1832e-10 |
| 5 | 5 | 0.01 | 6 | 7 | 6.6 | 0.047077 | 0.16567 | 0.058094 | 4.9033e-06 | 3.174e-12 |
| 5 | 1 | 0.1 | 9 | 10 | 9.91 | 0.075131 | 0.15946 | 0.095228 | 3.7638e-06 | 2.0527e-12 |
| 5 | $\frac{10}{(k+1)^2}$ | 0.1 | 10 | 12 | 11 | 0.077351 | 0.12221 | 0.095089 | 5.5684e-05 | 3.1871e-10 |
| 25 | 0.1 | 0 | 11 | 11 | 11 | 0.25215 | 0.30694 | 0.27212 | 5.0102e-05 | 5.0643e-11 |
| 25 | 5 | 0.01 | 5 | 6 | 5.98 | 0.090076 | 0.13349 | 0.10827 | 4.8911e-06 | 6.3716e-13 |
| 25 | 1 | 0.1 | 7 | 7 | 7 | 0.11813 | 0.16889 | 0.13914 | 4.5943e-06 | 6.0743e-13 |
| 25 | $\frac{10}{(k+1)^2}$ | 0.1 | 7 | 7 | 7 | 0.098635 | 0.14243 | 0.12109 | 3.5828e-06 | 4.4739e-13 |
| 50 | 0.1 | 0 | 8 | 8 | 8 | 0.57692 | 1.029 | 0.67213 | 8.1799e-05 | 6.7226e-11 |
| 50 | 5 | 0.01 | 5 | 5 | 5 | 0.24895 | 0.43507 | 0.34478 | 4.6578e-06 | 3.136e-12 |
| 50 | 1 | 0.1 | 7 | 7 | 7 | 0.35853 | 0.51106 | 0.43423 | 4.7164e-06 | 2.9162e-13 |
| 50 | $\frac{10}{(k+1)^2}$ | 0.1 | 6 | 7 | 6.99 | 0.2952 | 0.42262 | 0.36601 | 3.8439e-06 | 2.2537e-13 |
| 100 | 0.1 | 0 | 7 | 7 | 7 | 1.2405 | 1.828 | 1.4523 | 1.9767e-05 | 1.9599e-12 |
| 100 | 5 | 0.01 | 5 | 5 | 5 | 0.59282 | 0.95723 | 0.78223 | 4.2715e-06 | 1.333e-12 |
| 100 | 1 | 0.1 | 6 | 6 | 6 | 0.83413 | 1.2492 | 0.98612 | 4.8153e-06 | 1.563e-13 |
| 100 | $\frac{10}{(k+1)^2}$ | 0.1 | 6 | 6 | 6 | 0.62338 | 1.0542 | 0.81505 | 5.3529e-06 | 1.8962e-13 |

Tabela 4.2: Algoritmos MPP (SOUZA; OLIVEIRA, 2015) e MPPI para Problema 2.

| n | μ_k | γ_k | min. iter.(k) | max. iter.(k) | med. iter.(k) | min. time | max. time | med. time | grad f(X ^k) | f(X ^k) |
|-----|---------------------|------------|---------------|---------------|---------------|-----------|-----------|-----------|-------------------------|--------------------|
| 5 | 1 | 0 | 8 | 11 | 10.69 | 0.16139 | 0.3325 | 0.25947 | 1.2042e-04 | 1.9652e-10 |
| 5 | 0.1 | 0.01 | 11 | 14 | 13.76 | 0.1568 | 0.24653 | 0.22278 | 8.6164e-05 | 9.6829e-11 |
| 5 | 0.5 | 0.1 | 9 | 13 | 11.86 | 0.14281 | 0.24853 | 0.20879 | 5.4389e-05 | 3.9383e-11 |
| 5 | $\frac{2}{(k+1)^2}$ | 0.1 | 10 | 17 | 15.76 | 0.14788 | 0.26679 | 0.22738 | 8.4186e-04 | 9.1446e-09 |
| 25 | 1 | 0 | 11 | 11 | 11 | 0.94859 | 1.4397 | 1.1847 | 5.4563e-04 | 7.4563e-10 |
| 25 | 0.1 | 0.01 | 12 | 12 | 12 | 0.55303 | 0.71977 | 0.63063 | 3.4307e-04 | 2.9478e-10 |
| 25 | 0.5 | 0.1 | 12 | 12 | 12 | 0.69976 | 0.96456 | 0.81632 | 2.0127e-04 | 1.015e-10 |
| 25 | $\frac{2}{(k+1)^2}$ | 0.1 | 13 | 13 | 13 | 0.53209 | 0.91111 | 0.65692 | 4.4119e-04 | 4.8757e-10 |
| 50 | 1 | 0 | 11 | 11 | 11 | 3.0508 | 4.3938 | 3.3717 | 9.7601e-05 | 1.1912e-09 |
| 50 | 0.1 | 0.01 | 12 | 12 | 12 | 1.784 | 2.3953 | 1.9991 | 4.4926e-04 | 2.5242e-10 |
| 50 | 0.5 | 0.1 | 12 | 12 | 12 | 2.1716 | 3.2668 | 2.4341 | 3.429e-04 | 1.4707e-10 |
| 50 | $\frac{2}{(k+1)^2}$ | 0.1 | 12 | 12 | 12 | 1.5111 | 2.4 | 1.9059 | 8.12e-04 | 8.2456e-10 |
| 100 | 1 | 0 | 11 | 11 | 11 | 7.4373 | 11.1317 | 8.3472 | 1.6744e-03 | 1.7525e-09 |
| 100 | 0.1 | 0.01 | 12 | 12 | 12 | 4.1062 | 6.053 | 4.8483 | 6.5633e-04 | 2.693e-10 |
| 100 | 0.5 | 0.1 | 12 | 12 | 12 | 5.2686 | 8.4117 | 6.3045 | 5.7401e-04 | 2.0599e-10 |
| 100 | $\frac{2}{(k+1)^2}$ | 0.1 | 12 | 12 | 12 | 4.2017 | 6.2018 | 4.988 | 9.1262e-04 | 5.2065e-10 |

Tabela 4.3: Algoritmos MPP e MPPI para Problema 3.

| μ_k | γ_k | min. iter.(k) | max. iter.(k) | med. iter.(k) | med. iter.(k) | min. time(s) | max. time(s) | med. time(s) | med. time(s) | $\ C(x^k)\ $ |
|------------------------|------------|---------------|---------------|---------------|---------------|--------------|--------------|--------------|--------------|--------------|
| 0.1 | 0 | 187 | 259 | 244.15 | 244.15 | 0.0003684 | 0.001284 | 0.00070324 | 0.00070324 | 7.1621e-06 |
| 0.5 | 0.1 | 52 | 68 | 64.88 | 64.88 | 0.0001189 | 0.000402 | 0.00022221 | 0.00022221 | 1.2119e-06 |
| 1 | 0.5 | 22 | 29 | 27.7 | 27.7 | 5.19e-05 | 0.0001852 | 9.7068e-05 | 9.7068e-05 | 5.4863e-07 |
| $\frac{10^3}{(k+1)^2}$ | 0.5 | 41 | 44 | 42.86 | 42.86 | 8.99e-05 | 0.0003203 | 0.00014966 | 0.00014966 | 4.948e-08 |

Tabela 4.4: Algoritmos MPP (SUN et al., 2003) e MPPI para Problema 4.

| μ_k | γ_k | min. iter.(k) | max. iter.(k) | med. iter.(k) | min. time(s) | max. time(s) | med. time(s) | max. $\ C(x^k)\ $ |
|---------|------------|---------------|---------------|---------------|--------------|--------------|--------------|-------------------|
| 1 | 0 | 4 | 56 | 26.99 | 0.0025276 | 0.18057 | 0.017934 | 2.2505e-07 |
| 5 | 0.1 | 4 | 45 | 22.44 | 0.0021067 | 0.026851 | 0.01377 | 9.8168e-08 |
| 1 | 0.01 | 5 | 42 | 26.05 | 0.0026901 | 0.036193 | 0.015676 | 2.0754e-07 |

Em cada uma das figuras, do lado esquerdo apresentamos a variação de $\frac{d(x^{k+1}, x^k)}{\mu_k}$ e $\|d^k\|$ para o MMPI, uma vez que o critério de parada desse método é $d(x^{k+1}, x^k) < \varepsilon$ e $\|d^k\| < \varepsilon$. Para o MPP, apresentamos a variação de $\frac{d(x^{k+1}, x^k)}{\mu_k}$. Dessa forma ilustramos que as condições (A3) para MPP e (A4) para o MPPI são satisfeitas. As figuras do lado direito, mostram o comportamento da imagem do campo ao longo da sequência.

Os experimentos apresentados comprovam a eficácia do MPPI para resolver diferentes tipos de problemas. No Problema 1, o MPP coincide com método do ponto proximal proposto por (FERREIRA; OLIVEIRA, 2002), no Problema 2 coincide com o método do ponto proximal para funções DC proposto em (SOUZA; OLIVEIRA, 2015) e no Problema 4 coincide com método do ponto proximal proposto por (SUN et al., 2003). Conforme observamos nas tabelas, o MPPI supera o MPP em média de tempo, mesmo o algoritmo tendo um passo intermediário a mais que o MPP. Nos Problemas 1, 3 e 4, o MPPI também ganha do MPP em média de iteradas. Somente no Problema 2, os dois métodos tiveram performance similar.

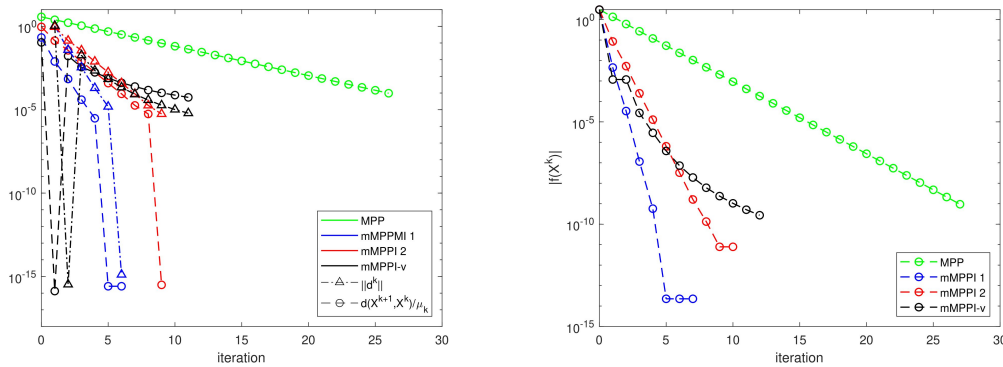


Figura 4.1: Algoritmos MPP e MPPI para Problema 1 com $n = 5$.

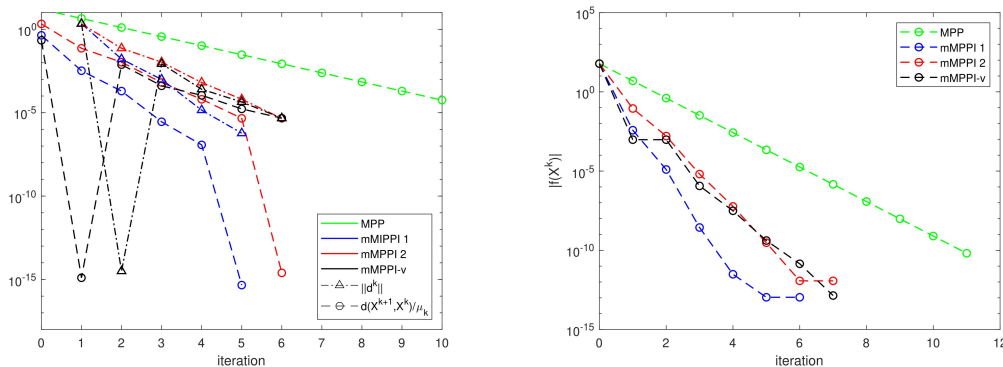


Figura 4.2: Algoritmos MPP e MPPI para Problema 1 com $n = 25$.

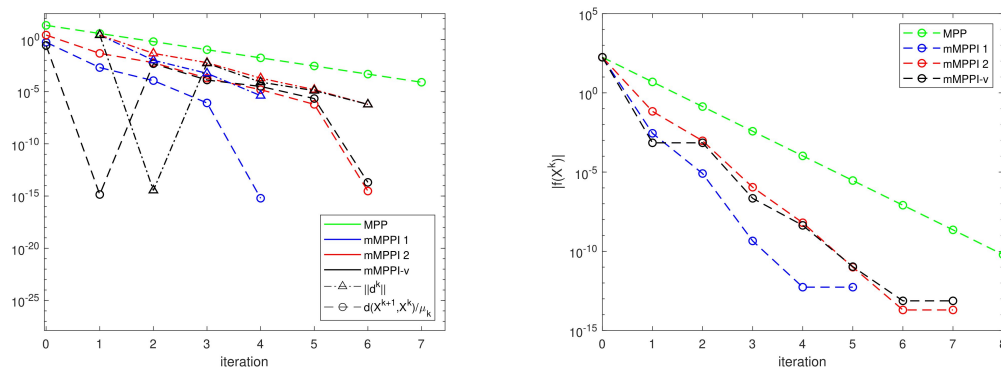


Figura 4.3: Algoritmos MPP e MPPI para Problema 1 com $n = 50$.

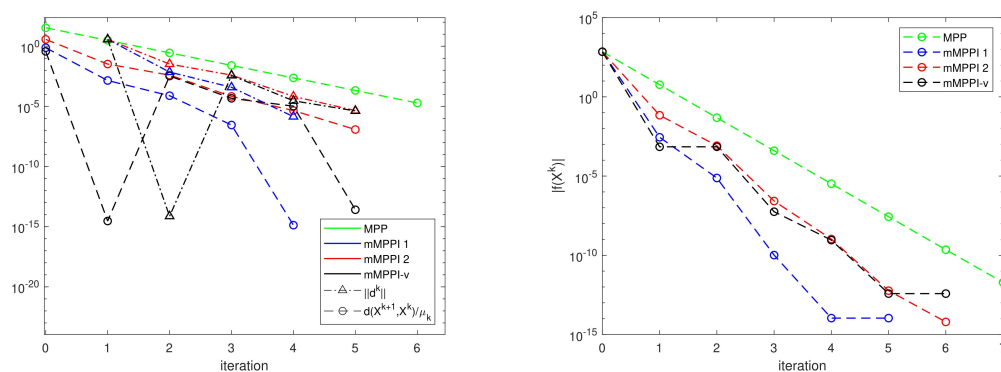


Figura 4.4: Algoritmos MPP e MPPI para Problema 1 com $n = 100$.

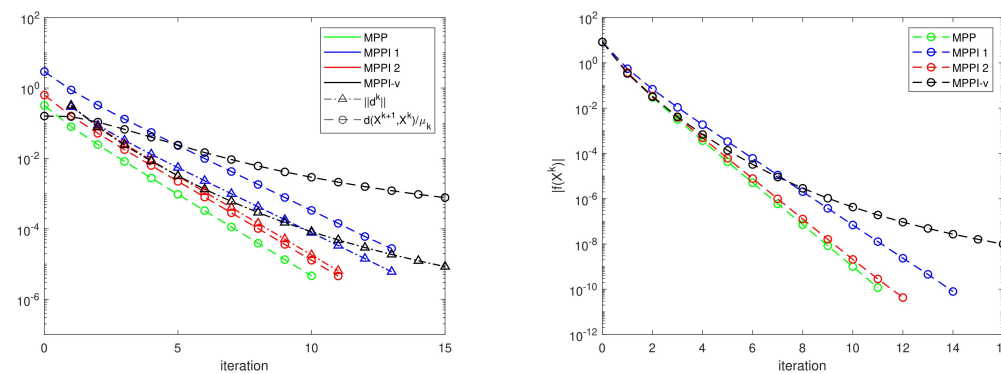


Figura 4.5: Algoritmos MPP e MPPI para Problema 2 com $n = 5$.

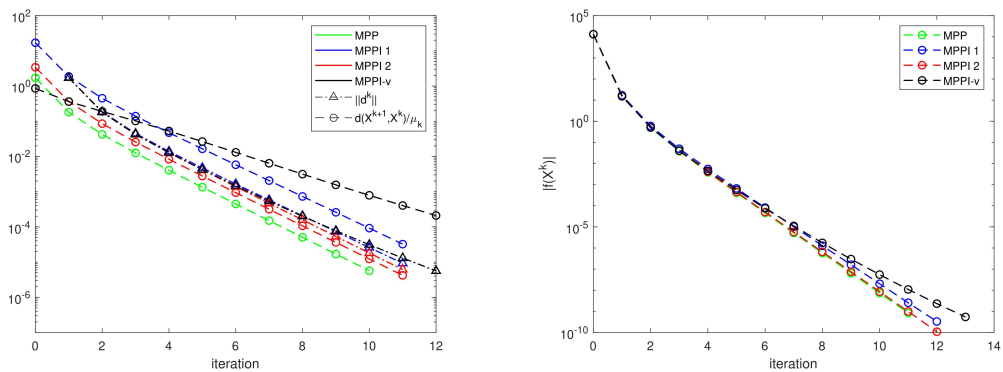


Figura 4.6: Algoritmos MPP e MPPI para Problema 2 com $n = 25$.

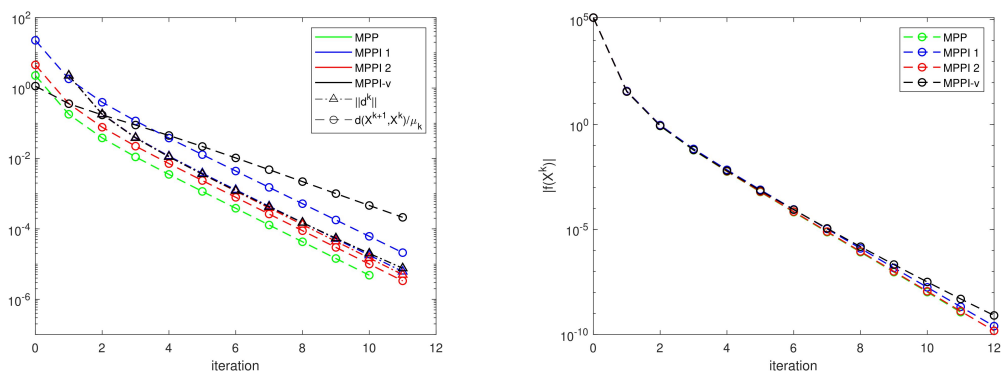


Figura 4.7: Algoritmos MPP e MPPI para Problema 2 com $n = 50$.

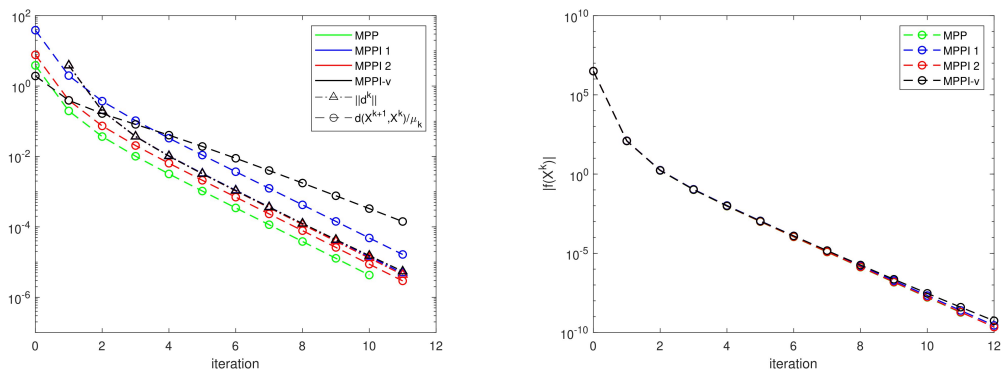


Figura 4.8: Algoritmos MPP e MPPI para Problema 2 com $n = 100$.

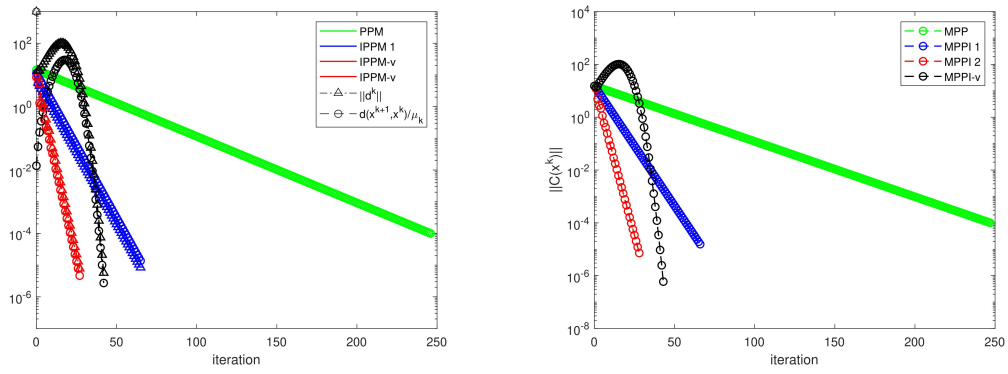


Figura 4.9: Algoritmos MPP e MPPI para Problema 3.

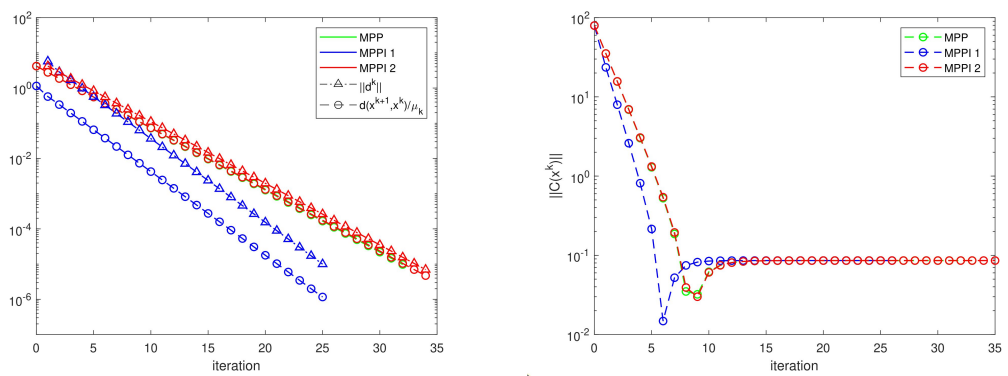


Figura 4.10: Algoritmos MPP e MPPI para Problema 4.

Capítulo 5

Considerações finais

Neste trabalho, apresentamos dois métodos do ponto proximal para encontrar zeros de campos de vetores possivelmente não-monótonos que são escritos como diferença de dois campos monótonos maximais em variedades de Hadamard. O primeiro método é uma extensão natural do método do ponto proximal para diferença de funções convexas e que generaliza o método do ponto proximal para campos vetoriais monótonos maximais. O segundo método é uma versão inercial do primeiro e é novo inclusive no contexto monótono em variedades de Hadamard. Para ambos os métodos, provamos os resultados padrões em programação DC mostrando que, caso exista, cada ponto de acumulação da sequência gerada pelos métodos é solução do problema. Além disso, sob hipóteses adicionais razoáveis, apresentamos algumas condições suficientes para limitação e convergência total dos métodos propostos. Tais condições são novas mesmo para o contexto das funções DC.

Ilustramos os métodos com alguns experimentos numéricos em uma variedade de Hadamard genuína (com curvatura não constante diferente de zero) e no espaço Euclidiano. Não era nossa intenção competir com métodos clássicos ou outros algoritmos especificamente desenvolvidos para o problema, mas sim mostrar que uma abordagem geral usando nossos métodos funciona bem em diferentes configurações, como problemas de minimização convexo e não convexo, inclusão monótona e não monótona. Acreditamos que um dos fatores que explica o interesse pelo método do ponto proximal está em sua simplicidade e facilidade de implementação para uma ampla classe de problemas. Por todas estas razões, surgiram várias variantes deste método e as suas propriedades foram descobertos ao longo dos anos, resultando em uma vasta literatura sobre o tema. Espera-

mos que nosso trabalho abra caminho para o desenvolvimento futuro de outros métodos para campos vetoriais não monótonos em variedades Riemannianas.

Acreditamos que uma possibilidade para pesquisas futuras seria propor um método inexato para os métodos apresentados neste trabalho. Um exemplo seria usar o conceito de alarguecimento de campos vetoriais em variedades de Hadamard como proposto em (BATISTA et al., 2016). Outra proposta seria estudar o caso restrito do problema abordado nesse trabalho.

Referências Bibliográficas

- ABSIL, P.-A.; MAHONY, R.; SEPULCHRE, R. Optimization algorithms on matrix manifolds. In: **Optimization Algorithms on Matrix Manifolds**. [S.l.]: Princeton University Press, 2009.
- ALIMOHAMMADY, M.; RAMAZANNEJAD, M. Inertial proximal algorithm for difference of two maximal monotone operators. **Indian Journal of Pure and Applied Mathematics**, Springer, v. 47, n. 1, p. 1–8, 2016.
- ALIMOHAMMADY, M.; RAMAZANNEJAD, M.; ROOHI, M. Notes on the difference of two monotone operators. **Optimization Letters**, Springer, v. 8, n. 1, p. 81–84, 2014.
- ALMEIDA, Y. T.; CRUZ NETO, J. X.; OLIVEIRA, P. R.; SOUZA, J. C. O. A modified proximal point method for DC functions on Hadamard manifolds. **Computational Optimization and Applications**, Springer, v. 76, n. 3, p. 649–673, 2020.
- ALVAREZ, F.; ATTOUCH, H. An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping. **Set-Valued Analysis**, Springer, v. 9, n. 1, p. 3–11, 2001.
- ANDRADE, J. S.; LOPES, J. O.; SOUZA, J. C. O. The proximal point method for difference of maximal monotone vector fields in Hadamard manifolds. **Submetido**, 2021.
- ANDRADE, J. S.; LOPES, J. O.; SOUZA, J. C. O. An inertial proximal point method for difference of maximal monotone vector fields in Hadamard manifolds. **Journal of Global Optimization**, p. 1–28, 2022.
- ANSARI, Q. H.; BABU, F. Existence and boundedness of solutions to inclusion problems for maximal monotone vector fields in Hadamard manifolds. **Optimization Letters**, Springer, v. 14, n. 3, p. 711–727, 2020.
- ANSARI, Q. H.; BABU, F. Proximal point algorithm for inclusion problems in Hadamard manifolds with applications. **Optimization Letters**, Springer, v. 15, n. 3, p. 901–921, 2021.
- ANSARI, Q. H.; BABU, F.; YA, J.-C. Inexact proximal point algorithms for inclusion problems on Hadamard manifolds. **JOURNAL OF NONLINEAR AND CONVEX ANALYSIS**, YOKOHAMA PUBL 101, 6-27 SATSUKIGAOKA AOBA-KU, YOKOHAMA, 227-0053, JAPAN, v. 21, n. 10, p. 2417–2432, 2020.
- ANSARI, Q. H.; BABU, F.; YAO, J.-C. Regularization of proximal point algorithms in Hadamard manifolds. **Journal of Fixed Point Theory and Applications**, Springer, v. 21, n. 1, p. 1–23, 2019.

- ARTACHO, F. J. A.; FLEMING, R. M.; VUONG, P. T. Accelerating the DC algorithm for smooth functions. **Mathematical Programming**, Springer, v. 169, n. 1, p. 95–118, 2018.
- ARTACHO, F. J. A.; VUONG, P. T. The boosted difference of convex functions algorithm for nonsmooth functions. **SIAM Journal on Optimization**, SIAM, v. 30, n. 1, p. 980–1006, 2020.
- ATTOUCH, H.; THÉRA, M. A general duality principle for the sum of two operators. **Journal of Convex Analysis**, Heldermann Verlag, v. 3, p. 1–24, 1996.
- BAČÁK, M. The proximal point algorithm in metric spaces. **Israel journal of mathematics**, Springer, v. 194, n. 2, p. 689–701, 2013.
- BAGIROV, A. M.; UGON, J. Codifferential method for minimizing nonsmooth DC functions. **Journal of Global Optimization**, Springer, v. 50, n. 1, p. 3–22, 2011.
- BATISTA, E. E. A.; BENTO, G. C.; FERREIRA, O. P. Enlargement of monotone vector fields and an inexact proximal point method for variational inequalities in Hadamard manifolds. **Journal of Optimization Theory and Applications**, Springer, v. 170, n. 3, p. 916–931, 2016.
- BATISTA, E. E. A.; BENTO, G. C.; FERREIRA, O. P. An extragradient-type algorithm for variational inequality on Hadamard manifolds. **ESAIM: Control, Optimisation and Calculus of Variations**, EDP Sciences, v. 26, p. 63, 2020.
- BENTO, G. C.; CRUZ NETO, J. X.; OLIVEIRA, P. R. A new approach to the proximal point method: convergence on general Riemannian manifolds. **Journal of Optimization Theory and Applications**, Springer, v. 168, n. 3, p. 743–755, 2016.
- BENTO, G. C.; FERREIRA, O. P.; MELO, J. G. Iteration-complexity of gradient, subgradient and proximal point methods on Riemannian manifolds. **Journal of Optimization Theory and Applications**, Springer, v. 173, n. 2, p. 548–562, 2017.
- BENTO, G. C.; FERREIRA, O. P.; OLIVEIRA, P. R. Proximal point method for a special class of nonconvex functions on Hadamard manifolds. **Optimization**, Taylor & Francis, v. 64, n. 2, p. 289–319, 2015.
- BENTO, G. C.; MELO, J. G. Subgradient method for convex feasibility on Riemannian manifolds. **Journal of Optimization Theory and Applications**, Springer, v. 152, n. 3, p. 773–785, 2012.
- BERGMANN, R.; HERZOG, R.; LOUZEIRO, M. S.; TENBRINCK, D.; VIDAL-NÚÑEZ, J. Fenchel duality theory and a primal-dual algorithm on Riemannian manifolds. **Foundations of Computational Mathematics**, Springer, v. 21, n. 6, p. 1465–1504, 2021.
- BHATIA, R. **Positive Definite Matrices**. [S.l.]: Princeton University Press, Princeton, 2009. v. 24.
- BORWEIN, J. M. Fifty years of maximal monotonicity. **Optimization Letters**, Springer, v. 4, n. 4, p. 473–490, 2010.

BOUMAL, N. An introduction to optimization on smooth manifolds. **Available online**, **May**, v. 3, 2022.

BOUMAL, N.; MISHRA, B.; ABSIL, P.-A.; SEPULCHRE, R. Manopt, a matlab toolbox for optimization on manifolds. **The Journal of Machine Learning Research**, JMLR. org, v. 15, n. 1, p. 1455–1459, 2014.

BREZIS, H. Operateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert. number 5 in north holland math. **Studies. North-Holland, Amsterdam**, Elsevier, 1973.

BREZIS, H.; HARAUX, A. Image d`une somme d`operateurs monotones et applications. **Israel Journal of Mathematics**, Springer, v. 23, n. 2, p. 165–186, 1976.

BROWDER, F. E. Fixed point theory and nonlinear problems. **Bulletin of the american mathematical society**, v. 9, n. 1, p. 1–39, 1983.

CRUZ NETO, J. X.; FERREIRA, O. P.; PÉREZ, L.; NÉMETH, S. Z. Convex-and monotone-transformable mathematical programming problems and a proximal-like point method. **Journal of Global Optimization**, Springer, v. 35, n. 1, p. 53–69, 2006.

CRUZ NETO, J. X.; FERREIRA, O. P.; PÉREZ, L. R. L. Monotone point-to-set vector field. **Balkan Journal of Geometry and Its Applications**, v. 5, n. 1, p. 69–79, 2000.

CRUZ NETO, J. X.; FERREIRA, O. P.; PÉREZ, L. R. L. Contributions to the study of monotone vector fields. **Acta Mathematica Hungarica**, Springer, v. 94, n. 4, p. 307–320, 2002.

CRUZ NETO, J. X.; LIMA, L. L. D.; OLIVEIRA, P. R. Geodesic algorithms in Riemannian geometry. **Balkan J. Geom. Appl**, v. 3, n. 2, p. 89–100, 1998.

CRUZ NETO, J. X.; OLIVEIRA, P. R.; SOUBEYRAN, A.; SOUZA, J. C. O. A generalized proximal linearized algorithm for DC functions with application to the optimal size of the firm problem. **Annals of Operations Research**, Springer, v. 289, n. 2, p. 313–339, 2020.

DO CARMO, M. P. **Geometria Riemanniana**. [S.l.]: Instituto de Matemática Pura e Aplicada(IMPA), Rio de Janeiro, 2015.

DOUGLAS, J.; RACHFORD, H. H. On the numerical solution of heat conduction problems in two and three space variables. **Transactions of the American mathematical Society**, JSTOR, v. 82, n. 2, p. 421–439, 1956.

FERREIRA, O. P.; LOUZEIRO, M. S.; PRUDENTE, L. Gradient method for optimization on Riemannian manifolds with lower bounded curvature. **SIAM Journal on Optimization**, SIAM, v. 29, n. 4, p. 2517–2541, 2019.

FERREIRA, O. P.; OLIVEIRA, P. R. Proximal point algorithm on Riemannian manifolds. **Optimization**, v. 51, n. 2, p. 257–270, 2002.

FERREIRA, O. P.; SANTOS, E. M.; SOUZA, J. C. O. Boosted scaled subgradient method for DC programming. **arXiv preprint arXiv:2103.10757**, 2021.

- HUANG, W.; WEI, K. Riemannian proximal gradient methods. **Mathematical programming**, Springer, v. 194, n. 1, p. 371–413, 2022.
- JOKI, K.; BAGIROV, A. M.; KARMITSA, N.; MAKELA, M. M.; TAHERI, S. Double bundle method for finding clarke stationary points in nonsmooth DC programming. **SIAM Journal on Optimization**, SIAM, v. 28, n. 2, p. 1892–1919, 2018.
- KHATIBZADEH, H.; RANJBAR, S. Inexact inertial proximal algorithm for maximal monotone operators. **Analele Universitatii”Ovidius” Constanta-Seria Matematica**, v. 23, n. 2, p. 133–146, 2015.
- LANG, S. **Fundamentals of differential geometry**. [S.l.]: Springer Science & Business Media, 2012. v. 191.
- LENGLET, C.; ROUSSON, M.; DERICHE, R.; FAUGERAS, O. Statistics on the manifold of multivariate normal distributions: Theory and application to diffusion tensor MRI processing. **Journal of Mathematical Imaging and Vision**, Springer, v. 25, n. 3, p. 423–444, 2006.
- LEUȘTEAN, L.; NICOLAE, A.; SIPOȘ, A. An abstract proximal point algorithm. **Journal of Global Optimization**, Springer, v. 72, n. 3, p. 553–577, 2018.
- LI, C.; LÓPEZ, G.; MARTÍN-MÁRQUEZ, V. Monotone vector fields and the proximal point algorithm on Hadamard manifolds. **Journal of the London Mathematical Society**, Oxford University Press, v. 79, n. 3, p. 663–683, 2009.
- LI, C.; LÓPEZ, G.; MARTÍN-MÁRQUEZ, V.; WANG, J.-H. Resolvents of set-valued monotone vector fields in Hadamard manifolds. **Set-Valued and Variational Analysis**, Springer, v. 19, n. 3, p. 361–383, 2011.
- MAINGÉ, P.-E.; MOUDAFI, A. Convergence of new inertial proximal methods for DC programming. **SIAM Journal on Optimization**, SIAM, v. 19, n. 1, p. 397–413, 2008.
- MARTINET, B. Regularisation, d’inéquations variationnelles par approximations succesives. **Revue Francaise d’informatique et de Recherche operationelle**, 1970.
- MINTY, G. J. Monotone networks. **Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences**, The Royal Society London, v. 257, n. 1289, p. 194–212, 1960.
- MOREAU, J. J. Proximité et dualité dans un espace hilbertien. **Bulletin de la Société mathématique de France**, v. 93, p. 273–299, 1965.
- MOUDAFI, A. On the difference of two maximal monotone operators: Regularization and algorithmic approaches. **Applied mathematics and computation**, Elsevier, v. 202, n. 2, p. 446–452, 2008.
- MOUDAFI, A. On critical points of the difference of two maximal monotone operators. **Afrika Matematika**, Springer, v. 26, n. 3, p. 457–463, 2015.
- MOUDAFI, A.; MAINGÉ, P.-E. On the convergence of an approximate proximal method for DC functions. **Journal of computational Mathematics**, JSTOR, p. 475–480, 2006.

- NÉMETH, S. Z. Monotone vector fields. **Publ. Math. Debrecen**, v. 54, n. 3-4, p. 437–449, 1999.
- NESTEROV, Y. E.; TODD, M. J. On the Riemannian geometry defined by self-concordant barriers and interior-point methods. **Foundations of Computational Mathematics**, Springer, v. 2, n. 4, p. 333–361, 2002.
- NOOR, M. A.; NOOR, K. I.; HAMDI, A.; EL-SHEMAS, E. H. On difference of two monotone operators. **Optimization Letters**, Springer, v. 3, n. 3, p. 329–335, 2009.
- OLIVEIRA, W. Proximal bundle methods for nonsmooth DC programming. **Journal of Global Optimization**, Springer, v. 75, n. 2, p. 523–563, 2019.
- OLIVEIRA, W. The ABC of DC programming. **Set-Valued and Variational Analysis**, Springer, v. 28, n. 4, p. 679–706, 2020.
- OLIVEIRA, W.; TCHEOU, M. P. An inertial algorithm for DC programming. **Set-Valued and Variational Analysis**, Springer, v. 27, n. 4, p. 895–919, 2019.
- PHAM, D. T.; SOUAD, E. B. Algorithms for solving a class of nonconvex optimization problems: Methods of subgradient. **FERMAT days 85: Mathematics for optimization**, Elsevier, North Holland Math. Stud., v. 129, p. 249–270, 1986.
- PHELPS, R. R. Lectures on maximal monotone operators. **Extracta Mathematicae**, Departamento de Matemáticas, v. 12, n. 3, p. 193–230, 1997.
- POLYAK, B. T. Some methods of speeding up the convergence of iteration methods. **Ussr computational mathematics and mathematical physics**, Elsevier, v. 4, n. 5, p. 1–17, 1964.
- POLYAK, B. T. Introduction to optimization. optimization software. **Inc., Publications Division, New York**, v. 1, p. 32, 1987.
- ROCKAFELLAR, R. Characterization of the subdifferentials of convex functions. **Pacific Journal of Mathematics**, Mathematical Sciences Publishers, v. 17, n. 3, p. 497–510, 1966.
- ROCKAFELLAR, R. On the maximal monotonicity of subdifferential mappings. **Pacific Journal of Mathematics**, Mathematical Sciences Publishers, v. 33, n. 1, p. 209–216, 1970.
- ROCKAFELLAR, R. T. Monotone operators and the proximal point algorithm. **SIAM journal on control and optimization**, SIAM, v. 14, n. 5, p. 877–898, 1976.
- ROTHAUS, O. S. Domains of positivity. In: **Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg**. [S.l.: s.n.], 1960. v. 24, n. 1, p. 189–235.
- SAKAI, T. **Riemannian Geometry**. [S.l.]: American Mathematical Soc., 1996. v. 149.
- SOUZA, J. C. O. Proximal point methods for lipschitz functions on Hadamard manifolds: scalar and vectorial cases. **Journal of Optimization Theory and Applications**, Springer, v. 179, n. 3, p. 745–760, 2018.

- SOUZA, J. C. O.; OLIVEIRA, P. R. A proximal point algorithm for DC functions on Hadamard manifolds. **Journal of Global Optimization**, Springer, v. 63, n. 4, p. 797–810, 2015.
- SOUZA, J. C. O.; OLIVEIRA, P. R.; SOUBEYRAN, A. Global convergence of a proximal linearized algorithm for difference of convex functions. **Optimization Letters**, Springer, v. 10, n. 7, p. 1529–1539, 2016.
- SUN, W.-y.; SAMPAIO, R. J.; CANDIDO, M. Proximal point algorithm for minimization of DC function. **Journal of computational Mathematics**, JSTOR, p. 451–462, 2003.
- TAO, P. D.; AN, L. T. H. Convex analysis approach to DC programming: theory, algorithms and applications. **Acta mathematica vietnamica**, v. 22, n. 1, p. 289–355, 1997.
- THI, H. A. L.; HUYNH, V. N.; DINH, T. P. Convergence analysis of difference-of-convex algorithm with subanalytic data. **Journal of Optimization Theory and Applications**, Springer, v. 179, n. 1, p. 103–126, 2018.
- THI, H. A. L.; PHAM, T. D. DC programming and DCA: thirty years of developments. **Mathematical Programming**, Springer, v. 169, n. 1, p. 5–68, 2018.
- UDRISTE, C. **Convex Functions and Optimization Methods on Riemannian Manifolds, Mathematics and its Applications**. [S.l.]: Kluwer Academic Publishers, Dordrecht, Netherlands, 1994.
- WANG, J.; LI, C.; LOPEZ, G.; YAO, J.-C. Convergence analysis of inexact proximal point algorithms on Hadamard manifolds. **Journal of Global Optimization**, Springer, v. 61, n. 3, p. 553–573, 2015.
- WANG, J.; LI, C.; LOPEZ, G.; YAO, J.-C. Proximal point algorithms on Hadamard manifolds: linear convergence and finite termination. **SIAM Journal on Optimization**, SIAM, v. 26, n. 4, p. 2696–2729, 2016.